To Read or Not to Read

Suzanne E. Mol

Printed by Mostert & Van Onderen, Leiden

To Read or Not to Read

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 7 december 2010
klokke 15:00 uur

door

Suzanne Elizabeth Mol

geboren te Naarden
in 1983

Promotiecommissie

Promotor:
Prof. dr. A. G. Bus

Overige leden:
Prof. dr. M. H. van IJzendoorn
Prof. dr. P. W. van den Broek
dr. P. F. de Jong (*Universiteit van Amsterdam*)

ISED
Institute for the Study of Education and Human Development

Table of Contents

1

General Introduction

Parts of this chapter are based on:
Bus, A. G., Van IJzendoorn, M. H., & Mol, S. E. (in press). Meta-Analysis. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methods, volume 2*. New York: Guilford Press.

Entering search terms as “storybook reading” or “shared book reading” into electronic databases such as PsycInfo, ERIC, and ProQuest result in more than 300 hits of peer-reviewed journal articles, book chapters, and dissertations. At least another 100 studies focus on “leisure time reading” or “recreational reading” in conventionally reading students. In one way or another, this ever increasing set of correlational, longitudinal, and experimental studies tries to operationalize, prove, and/or nuance the appealing assumption that reading (story)books has a long-lasting impact on our language and reading abilities as well as our academic success. Do quantitative integrations of the research base thus far corroborate with the general belief in society that reading is as a miracle drug for the prevention and treatment of reading problems?

Meta-analysis is the empirical analysis of empirical studies – that is, the quantitative analysis and synthesis of a set of related empirical studies in a well-defined domain (Cooper, Hedges, & Valentine, 2009). The three meta-analyses in this thesis comprise 146 studies (N = 10,308) that address the role of book reading in language and reading development from infancy to early adulthood. For the group of pre-conventional readers, we also examine the effect of interventions that improve the quality of shared book reading such as dialogic reading programs (e.g., parents are trained to ask questions about words and story events; Whitehurst et al., 1988). Hereafter, I first introduce the importance of developing a book reading routine. Second, I elaborate on the meta-analytic approach we applied. In the third and final section, I present the aims and outline of this thesis.

The Quantity and Quality of Book Reading

We view parent-child book sharing as part of a continuum of leisure-time reading experiences that facilitate and influence language and reading skills throughout development (see chapter 2). Developing a book reading routine before the age of two may set in motion a causal spiral, in which language skills develop as a result of shared book reading and in which children’s vocabulary size determines whether they comprehend storybooks and whether they enjoy being read to (Fletcher & Reese, 2005). For conventional readers, this spiral continues to determine their reading behavior: Reading books is seen as both a consequence of reading ability and a contributor to further reading development (Stanovich, 2000). More skilled readers are more likely to choose to read more frequently which, in turn, will improve their knowledge of word forms and semantics, and enhance their vocabulary size and text comprehension abilities. In contrast, readers with small vocabularies or word reading difficulties may not succeed in comprehending text, become less eager to read, and as a result, show stagnation in their reading development (Kush, Watkins, & Brookhart, 2005). Because of growing individual differences in leisure-time reading activities, we expect that the relationship between book reading and reading skills will strengthen from infancy to early adulthood as frequent readers will be more motivated and better skilled readers than infrequent book readers.

Young children need caregivers who bridge the gap between the world of the book and their own world so reading storybooks is not only a source of entertainment but also is a means to get familiar with the structure and syntax of written language and learn about the purpose and function of reading (Bus, 2003; DeTemple & Snow, 2003; Heath, 1982; Scheele, Leseman, & Mayo, 2010; Sulzby, 1985; Watson, 2001). The richness of children's learning experiences is thought to depend not only on the frequency of book sharing but also on the quality of reading sessions and parents' sensitivity towards children's cognitive abilities and interests: More new words are learned when young children are actively involved in storybook reading, for instance when questions are asked about pictures, difficult words, and story events, and informative feedback is provided on children's answers (e.g., Whitehurst et al., 1988). Because observational studies have suggested that most parents – and especially parents in low socioeconomic status groups – do not apply such interactive reading techniques spontaneously (e.g., Britto, Brooks-Gunn, & Griffin, 2006; Bus & Van IJzendoorn, 1995; Heath, 1982; Laakso, Poikkeus, & Lyytinen, 1999; Ninio, 1980; Silvén, Ahtola, & Niemi, 2003), attempts are made to train caregivers in techniques such as "Dialogic Reading" that may enhance their quality of reading interactions. We integrate effects of training studies to estimate overall effects of such interventions as well as effects of age and risk status (see chapter 3).

At the start of formal schooling, children from low socioeconomic status homes are more likely to lag behind in their language and reading skills than their peers with a middle- or high socioeconomic status (Dickinson & McCabe, 2001; Hart & Risley, 2003). As such an achievement gap is likely to widen in the course of primary school (Alexander, Entwisle, & Olson, 2007; Foster & Miller, 2003), it seems especially important for children who are at risk for language and literacy impairments to have frequent encounters with storybooks from an early age onwards. However, because children with a low socioeconomic status are more likely to grow up in home environments that are less stimulating (e.g., parents read less frequently, own few (children's) books) than the home environments of their peers from higher socioeconomic backgrounds (Christian, Morrison, & Bryant, 1998; Morrow, 1983; Van Steensel, 2006), the quantity and quality of storybook reading in their preschool and kindergarten classrooms is deemed essential for enhancing the language and reading skills they need to benefit from formal reading instruction in primary school. This has resulted in interactive reading techniques as a means for teachers to attract the attention of at-risk children, to stimulate them to use book-related language, and to check their story understanding. Chapter 4 presents an integration of studies that test effects of interactive reading techniques applied in classroom settings.

Steps in a Meta-Analysis

Meta-analysis can be applied most fruitfully within research programs in which studies with similar designs or measures accumulate over the years. In primary studies, data are collected to test a hypothesis derived from a well-articulated theory; the hypothesis often will be stated in the form: variable *X* is associated with variable *Y*, or *X* is causally related to *Y*. In correlational or experimental designs, measures prototypical to assess *X* and *Y* are being used, and the results are therefore comparable across studies. If the number of replications increases, and if characteristics of replication studies vary, the meta-analytic approach is feasible to synthesize the literature and to test the effects of variations in study characteristics on the outcome of the studies.

A common and defining characteristic of all meta-analytic approaches is the use of a specific set of statistical methods compared to the methods used in primary research. The reason is simple: In primary research the unit of analysis is the individual participant (or class, or other group), whereas the unit of meta-analysis is the study result. Study results are usually based on different numbers of participants, and they are, therefore, point estimates with different precision and confidence boundaries (Mullen, 1989). It would be incorrect to give a significant correlation of .30 in a sample of 50 participants (confidence interval: .02, .53) the same weight as a correlation of .30 in a sample of 500 participants (confidence interval: .22, .38). Basically, however, meta-analytic research follows the same steps and standards as empirical research (Cooper, 1982).

Step 1: Hypothesis Formulation

The meta-analysis should start with the formulation of a specific, theoretically relevant conceptual framework. Its domain should be clearly defined, and the central meta-analytic question should be theoretically derived and meaningful. When a meta-analyst is not sensitive to such substantive issues, a meta-analysis can become a pointless, merely statistical exercise (Littell, Corcoran, & Pillai, 2008). For example, when synthesizing the effects of interventions on struggling readers' reading comprehension (see Edmonds et al., 2009), the validity and/or practical use of the summary effects can be questioned when interventions with a focus on fluency, decoding, comprehension, and multiple components are heaped together. That is, even though the dependent measure is comparable and the target groups are similar, it is difficult to disentangle the kind of intervention that might support the comprehension skills of children with reading disabilities when the content of interventions varies extensively.

Step 2: Retrieval and Coding of Studies

In the next stage, the meta-analyst should systematically collect the relevant published as well as unpublished literature from at least three different sources.

The "snowball" method (using references lists from key articles in the field), the "invisible college" approach (using key figures in the field to collect recent or unpublished materials), and computer searches of subject indexes such as ERIC, PsycInfo, Medline, Proquest UMI Dissertations, and Google Scholar, or citation indexes such as SSCI or SCI may be used in a multimethod combination.

In some meta-analytic approaches, selection of studies is based on the idea that only randomized experimental designs produce valid findings to be taken seriously. "What Works Clearinghouse" (WWC; see http://ies.ed.gov/ncee/wwc/), that is set up to support educators and the U.S. Department of Education in making evidence-based recommendations about the effectiveness of programs, policies, and practices in a wide range of areas, uses eligibility screens in which randomization, level of attrition, and equivalence of treatment and control groups are taken into account to select studies that meet the evidence standards fully or with reservations (WWC, 2008). The National Reading Panel (NRP) also objects to inclusion of all studies regardless of design features. Restrictions of the type of papers to be included, however, may imply an untenable reduction of the available evidence. For instance, the NRP discards the many correlational investigations in the area of reading research (NRP, 2000; Williams, 2001) which means a loss of potentially important information. In this respect, the meta-analytic method is basically indifferent: The central hypothesis should decide about the feasibility of selection criteria, and when this hypothesis is not stated in strictly causal terms there is no reason to leave correlational studies aside. Furthermore, the impact of the quality of research on effect sizes can be examined by testing whether the overall effect is influenced by the presence of studies with other designs than randomized controlled trials (Rosenthal, 1995).

The exhaustive search for pertinent literature is preferred compared with the best evidence approach (Slavin, 1986), in which only the qualitatively sound studies would be allowed to enter a meta-analysis. Because of their emphasis on explanation of variability in effect sizes, in recent meta-analytic approaches it is preferred to test whether quality of research (which always is a matter of degree and a matter of different strengths and weaknesses) explains variation in study results in order to make the process of study selection and evaluation transparent and to maximize the power of the analyses. We, for instance, created a scale to score whether the researchers checked the use of trained techniques in the experimental group, the quality of reading sessions within the control group, and the actual frequency of book reading in the experimental and the control groups (see chapter 4). Experimental designs outperformed quasi-experiments on the scale, but intervention outcomes were not affected by the experiment fidelity score nor did quasi-experiments reveal higher effect sizes than true experiments for children's language and literacy measures.

The basic problem to be faced in the second stage of the meta-analysis is the "file-drawer" problem (Rosenthal, 1991). Primary researchers know that it is easier to get papers published in which they report significant results than to guide papers into print with null results; regardless of the quality of the study (Begg, 1994). This publication bias may even lead to the unfortunate situation that the majority of papers remain in the file drawers of disappointed researchers, whereas only a minority of papers with significant results are published (Cohen, 1990). Average or combined effect sizes of published papers may, therefore, present an inflated picture of the real state of the art in a specific field. The number of unpublished papers with null findings that are needed to make the meta-analytic outcome insignificant can be estimated (the "fail-safe number"; Rosenthal, 1991). A publication bias can be visually inspected by a funnel plot, which is a scatter plot of the effect size against sample size that will be skewed and show asymmetry (i.e., due to a lack of small effect sizes) when a publication bias is present (Lipsey & Wilson, 2001). The "file-drawer" problem may suggest that *a priori* selection of only published papers is not always warranted. Although published studies have been subjected to more or less thorough reviewing procedures and therefore seem to carry more quality weight than unpublished studies, the reasons for remaining unpublished may be unrelated to quality. In many cases, it is, therefore, better to collect all studies regardless of origin or status, and to analyze *post hoc* whether publication status makes a difference in combining effect sizes. To assess the likely impact of a publication bias, the "trim and fill" method can be used to estimate the unbiased effect size, by estimating the number of missing studies from an asymmetrical funnel plot and calculating an adjusted point estimate and variance (Duval & Tweedie, 2000a, 2000b).

The retrieved articles, dissertations, and unpublished documents are considered to be the raw data to which a coding system is applied to produce the variables to be used in the meta-analysis. The application of the coding system should be tested for intercoder reliability. The coding system contains potential moderator variables that can be used to explain the variability of the effect sizes in the specific set of studies. The variables in the coding system should therefore be theoretically relevant and constitute pertinent moderator hypotheses. In view of the relatively small number of studies included in most meta-analyses, the coding system should not be too extended. If potential moderators exceed the number of studies, inflated meta-analytic outcomes may be the nonreplicable result. On the other hand, if the number of studies per moderator is too small, the power will be too low to detect meaningful differences in effects across subgroups (Hedges & Pigott, 2004). We suggest that a minimum of four (Bakermans-Kranenburg, van IJzendoorn, & Juffer, 2003) or five (Seidel & Shavelson, 2007) studies per subgroup is needed to reliably interpret any contrast.

Studies may report effects on several dependent measures for similar outcome measures. To avoid a situation in which studies with more results have a greater impact, the effect sizes should be aggregated within studies and domains. For example, when a study reports outcomes for one receptive and two expressive vocabulary measures, the expressive outcomes are aggregated in a first step (domain) and combined with the effect size of the single receptive measure in a second step in order to calculate a vocabulary composite per study. Separate meta-analyses can be conducted to examine differential treatment effects per outcome measure (e.g., Does dialogic reading affect expressive vocabulary more strongly than receptive vocabulary skills?; see chapter 3) as long as each study contributes one effect size to each analysis. Creating a rather broad composite such as "academic achievement", in which a variety of reading, mathematical, and/or grade-related measures are aggregated, might limit the interpretation of specific treatment effects. Another complicating factor may be that experiments include two or more interventions but only one control group. Effect sizes are dependent if the same control group is used to calculate the effect sizes for each treatment (e.g., Ehri, Nunes, Willows, et al., 2001). Gleser and Olkin (1994) state that in multiple-treatment studies, "the treatments may all be regarded as instances or aspects of a common treatment construct." Furthermore they state that "there is strong reason a priori to believe that a composite effect size of treatment obtained by combining the end point effect sizes would adequately summarize the effect of treatment" (p. 351). Another, more pragmatic solution of the multiple-interventions problem is to divide the sample size of the control group in the same number of subgroups as there are interventions in order to avoid the situation in which control subjects count for more than one unit of analysis.

Step 3: Analysis of Study Results and Characteristics

Data analysis often consists of three steps (Mullen, 1989): First, the central tendency of the study results is computed (i.e., the combined effect size). Because *p* values heavily depend on the number of observations, recent meta-analyses focus on the combined standardized differences between the means of the experimental and the control groups. The statistic used to assess the effectiveness of a treatment or other variable is the effect size, *d*, which measures how much the mean of the treatment group exceeds the mean of the control group in standard deviation units. Effect size expresses how many standard deviation units treatment groups differ from control groups without treatment. An effect size of 1.0 indicates that the treatment group mean is one standard deviation higher than the control group mean, whereas an effect size of 0 indicates that treatment and control group means are identical. A mean effect, of which the precision is addressed by the 95% confidence interval (CI), is considered significant if the CI does not include zero. Differences between estimates can be interpreted as significant when the CIs do not overlap. According to Cohen (1988), an effect size of $d = .20$ is considered small, an effect size of $d = .50$ moderate, and an effect size of $d = .80$ or above large ($r = .10$ is small, $r = .30$ is moderate, $r = .50$ is large). Translated into percentiles, $d = .20$ indicates that the treatment has moved the average child from the 50th to the 58th percentile; $d = .50$ indicates that the treatment has moved the child, on average, to the 69th percentile; d = .80 indicates that the treatment has moved the child, on average, to the 79th percentile. As an alternative, Rosenthal and Rubin (1982) suggested the binominal effect size display (BESD), which indicates the change in predictive accuracy attributable to the relationship in question and is computed from the formula .50 ± (r/2). The BESD shows the extent to which prediction is enhanced (i.e., the percentage increase in prediction) with the use of intervention X to predict reading skill Y (for details, see step 4).

A weighted effect size is mostly used to adjust for the bias resulting from small sample sizes (i.e., the tendency of studies with small samples to overestimate effects). Unweighted *d*'s are sometimes presented to provide information about the direction of biases related to sample size. The effect size can be computed on the basis of the standard deviations of the control group (Glass, 1976), the pooled standard deviations (Rosenthal, 1991), or the pooled variance (Hedges & Olkin, 1985). Cohen's d, for instance, is calculated as the difference between control and experimental treatment posttest mean scores (partialed for the influence of pretest scores if information is available) divided by the pooled standard deviation. Alternatively, the test statistics (F, t, χ^2) can be transformed into an effect size (Rosenthal, 1991). In practice, different strategies do not seem to make a substantial difference (Johnson, Mullen, & Salas, 1995).

Second, the variability of the results around this central tendency is assessed, and outliers as well as homogeneous subsets of studies are identified. To determine whether a set of *d*'s shares a common effect size, a homogeneity statistic (Q) which approximates chi-square distribution with k – 1 degrees of freedom, where k is the number of effect sizes, can be computed. Homogeneity analysis compares the amount of variance exhibited by a set of effect sizes with the amount of variance expected if only sampling error is operating. I-squared (I^2), another indicator of homogeneity, describes the impact of heterogeneity on a meta-analysis by measuring the degree of inconsistency between studies. Values that exceed 70% should invite caution about the homogeneity of the mean effect (Petticrew & Roberts, 2006). If sets of study results remain heterogeneous, combined effect size computed on the basis of the fixed model may be biased estimates, that is, it cannot be concluded that they are a sample from the same population, and a random model should be preferred (Hedges, 1994). In the random-effects model, studies are also weighted by the inverse of its variance, but, in addition, it accounts for within-study error as well as between-study variation in true effects (Borenstein, Hedges, Higgins, & Rothstein, 2009). If a distribution of study results is extremely skewed and shows several outlying values, the average effect size does not adequately represent the central tendency. Inflated meta-analytic findings may result from ignoring heterogeneity in study outcomes, and the (more conservative) random model may lead to lower estimates for the combined effect size as well as larger confidence boundaries (Hedges, 1994).

Third, through a moderator analysis, the meta-analysts try to explain the variability on the basis of study characteristics. A significant chi-square indicates that the study features significantly moderate the magnitude of effect sizes. For example, intervention studies with randomized designs may, on average, yield smaller effects than those without randomization. Mostly the analyses do not include tests of interactions between moderator variables because the number of comparisons is insufficient in many cases. It should be noted that in meta-analytic as well as in primary studies every subject or sample should be counted independently from each other and only once. That is, if a study presents more than one effect size for the same hypothesis, these effect sizes should be combined within the study before it is included in the overall meta-analysis.

Step 4: Interpretation of Meta-Analytic Outcomes

The interpretation of the size of the combined effects is a matter of much debate (McCartney & Rosenthal, 2000). In a meta-analysis, Van IJzendoorn and Bus (1994) showed that a powerful explanation of dyslexia, the phonological deficit hypothesis, explains only 6% of the variance in dyslexia (d = .48) which is about half a standard deviation difference between the experimental and the control groups. Bus, van IJzendoorn, and Pellegrini (1995) showed that the association between preschool storybook sharing and later literacy was even stronger (d = .59) explaining about 8% of the variance in children's literacy skills. A correlation of .28 between book sharing and reading may seem a rather modest outcome. However, in terms of the BESD (Rosenthal, 1991), this effect is sizable. The BESD is defined as the change in success ratio because of an intervention. The BESD shows the extent to which prediction is enhanced (i.e., the percentage increase in prediction) with the use of intervention X to predict reading skill Y. If we equal the combined effect size d = .59 with an r = .28, the success ratio in the experimental group would be: .50 + (.28/2) = .64; the success ratio in the control group would be .50 – (.28/2) = .36. It should be noted, therefore, that it certainly can make a tremendous difference in the lives of young children whether or not they are read to by their parent. The difference between the experimental and the control groups would amount to a substantial difference if we translate this outcome to the millions of children who may profit from book reading (Rosenthal, 1991). Taking into account that experimental studies revealed outcomes similar to correlational/longitudinal/retrospective studies, this meta-analysis provides a clear and affirmative answer to the question of whether or not storybook reading is one of the most important activities for developing the knowledge required for eventual success in reading. Therefore, parental storybook reading should be recommended because in terms of BESD it makes a difference for many thousands of preschoolers. In the same vein, phonological deficit is correctly considered as a main cause of dyslexia.

Aims and Outline Thesis

A previous meta-analysis (Bus et al., 1995), comprising 33 studies between 1951 and 1993, showed that reading storybooks to pre-conventional reading children explained about 10-12% of children's oral language skills (i.e., passively comprehending words and/or actively producing words) and 8% of the variance in children's basic reading skills such as knowledge of letter names, how letters relate to sounds in spoken words, and how to write your own name. The meta-analyses in this thesis do not only include a more recent set of studies than in Bus et al. (1995), but also extend the age range from infancy to early adulthood and examine more closely the impact of qualitative aspects of shared book reading. The latter meta-analyses are a critical test of early interventions that are designed to improve pre-conventional readers' literacy experiences at home and at school. Among the questions that guided our attempt to synthesize the available literature were:

a) Does exposure to (story)books affect language proficiency and does it get even stronger in adulthood?
b) Does book sharing already stimulate technical reading skills in pre-conventional readers and do these skills improve beyond the earliest stages of reading acquisition?
c) Do students' ability to spell words correctly depend on exposure to print?
d) Is it advisable to promote reading outside school in low-ability readers?
e) Does a reading habit enhance intellectual abilities and later success in society?
f) Do young children whose caregivers (i.e., parents, teachers) ask questions about words and story events during shared book reading learn more words from storybooks than children who are only read the story text?

In **chapter 2**, we relate leisure-time reading activities of (1) preschoolers and kindergartners, (2) children in grades 1 to 12, and (3) college and university students to indicators of comprehension and technical reading and spelling skills. We examine (a) whether the pattern of associations gets stronger across the age span from early childhood to early adulthood, and (b) the extent to which low-ability readers benefit from independent book reading.

In **chapter 3**, we meta-analyze the effects of Dialogic Reading interventions for parents of 2- to 6-year-old children. We expected that children's expressive vocabulary skills (i.e., producing words) are especially affected, because those skills are particularly emphasized within the Dialogic Reading format. Furthermore, we tested (a) whether children at risk for language and literacy impairments (e.g., due to a low socioeconomic status) benefited less from the intervention, because their parents may be less responsive to training in book-sharing skills, and (b) whether older children with more linguistic skills were affected less than younger children, because older children may be less dependent on book reading quality.

In **chapter 4**, we quantitatively summarize (quasi-)experiments that examined the effects of interactive reading in preschool and kindergarten classrooms. We expected that (a) children's oral language skills would improve as a result of the intervention, and (b) basic reading skills would be especially affected in older, kindergarten children who may be capable of interacting with the adult and simultaneously process other features of the printed text. We also examined whether intervention success was moderated by children's risk status, the size of the interactive reading groups, and whether the implementer was a researcher or children's own teacher.

In the **chapter 5**, a meta-analysis of meta-analyses will be presented and related to questions for future research.

References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72,* 167-180.

Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129,* 195-215.

Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 399-409). New York: Russell Sage Foundation.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* West Sussex, U.K.: Wiley.

Britto, P. R., Brooks-Gunn, J., & Griffin, T. M. (2006). Maternal reading and teaching patterns: Associations with school readiness in low-income African American families. *Reading Research Quarterly, 41,* 68-89.

Bus, A. G. (2003). Social-emotional requisites for learning to read. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 3-15). Mahwah, NJ: Lawrence Erlbaum Associates.

Bus, A. G., & van IJzendoorn, M. H. (1995). Mothers reading to their three-year-olds: The role of mother-child attachment security in becoming literate. *Reading Research Quarterly, 30*, 998-1015.

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1–21.

Christian, K., Morrison, F. J., & Bryant, F. B. (1998). Predicting kindergarten academic skills: Interactions among child care, maternal education, and family literacy environments. *Early Childhood Research Quarterly, 13,* 501-521.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45,* 1304–1312.

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research, 52*, 291–302.

DeTemple, J., & Snow, C.E. (2003). Learning words from books. In A. van Kleeck, S. A. Stahl & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 16-36). Mahwah, NJ: Lawrence Erlbaum.

Dickinson, D. K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research & Practice, 16*, 186-202.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method for accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89-98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455-463.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K. et al. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79,* 262-300.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36,* 250–287.

Fletcher, K. L., & Reese, E. (2005). Picture book reading with young children: A conceptual framework. *Developmental Review, 25,* 64-103.

Foster, W.A., & Miller, M. (2007). Development of the literacy achievement gap: A longitudinal study of kindergarten through third grade. *Language, Speech, and Hearing Services in Schools, 38,* 173-181.

Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Research, 5*, 3–8.

Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 339–355). New York: Russell Sage.

Hart, B., & Risley, T. (2003). The early catastrophe: The 30 million word gap. *American Educator, 27,* 4-9.

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11,* 49-76.

Hedges, L. V. (1994). Statistical considerations. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 29–38). New York: Russell Sage.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9,* 426-445.

Johnson, B. T., Mullen, B., & Salas, E. (1995). Comparison of three meta-analytic approaches. *Journal of Applied Psychology*, *80*, 94–106.

Kush, J. C., Watkins, M. W., & Brookhart, S. M. (2005). The temporal-interactive influence of reading achievement and reading attitude. *Educational Research and Evaluation, 11,* 29-44.

Laakso, M. L., Poikkeus, A. M., &Lyytinen, H. (1999). Shared reading interaction in families with and without genetic risk for dyslexia: Implications for toddlers' language development. *Infant and Child Development, 8,* 179-195.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Littell, J. H., Corcoran, J., & Pillai, V. K. (2008). *Systematic reviews and meta-analysis.* New York: Oxford University Press.

McCartney, K., & Rosenthal, R. (2003). Effect Size, practical importance, and social policy for children. *Child Development, 71,* 173–180.

Morrow, L. M. (1983). Home and school correlates of early interest in literature. *Journal of Educational Research, 76,* 221-230.

Mullen, B. (1989). *Advanced basic meta-analysis*. Hillsdale, NJ: Erlbaum.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Rockville, MD: NICHD Clearinghouse.

Ninio, A. (1980). Picture-book reading in mother-infant dyads belonging to two subgroups in Israel. *Child Development, 51,* 587-590.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences; A practical guide.* Oxford: Blackwell Publishing.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (rev. ed.). Newbury Park, CA: Sage.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183-192.

Rosenthal, R., & Rubin, D. B. (1982). Further meta-analytic procedures for assessing cognitive gender differences. *Journal of Educational Psychology, 74*, 708–712.

Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics, 31*, 117-140.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77,* 454-499.

Silvén, M., Ahtola, A., & Niemi, P. (2002). Early words, multiword utterances and maternal reading strategies as predictors of mastering word inflections in Finnish. *Journal of Child Language, 30,* 253-279.

Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15*, 5–11.

Stanovich, K. E. (2000). *Progress in understanding reading; Scientific foundations and new frontiers.* New York: The Guilford Press.

Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. *Reading Research Quarterly, 20,* 458-481.

Van IJzendoorn, M. H., & Bus, A. G. (1994). Meta-analytic confirmation of the nonword reading deficit in developmental dyslexia. *Reading Research Quarterly, 29,* 266–275.

Van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. *Journal of Research in Reading, 29*, 367-382.

Watson, R. (2001). Literacy and oral language: Implications for early literacy acquisition. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 43-53). New York: Guilford Press.

What Works Clearinghouse (2008, December). *Procedures and standards handbook (version 2.0).* Retrieved November 28, 2009, from http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., & Fischel, J. E. et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24,* 552-559. Doi: 10.1037/0012-1649.24.4.552

Williams, J. (2001). Commentary: Four meta-analyses and some general observations. *The Elementary School Journal, 101,* 349–354.

2

To Read or Not to Read:
A Meta-Analysis of Print Exposure
from Infancy to Early Adulthood

Abstract

This research synthesis examines whether the association between print exposure and components of reading grows stronger across development. We meta-analyzed 99 studies (total N = 7,669) that focused on leisure-time reading of (a) preschoolers and kindergartners, (b) children attending grade 1 to 12, and (c) college and university students. For all measures in the outcome domains of reading comprehension and technical reading and spelling, moderate to strong correlations with print exposure were found. The outcomes support an upward spiral of causality: Children who are more proficient in comprehension and technical reading and spelling skills read more; because of more print exposure their comprehension and technical reading and spelling skills improved more with each year of education. For example, in preschool and kindergarten print exposure explained 12% of the variance in oral language skills, in primary school 13%, in middle school 19%, in high school 30%, and in college and university 34%. Moderate associations of print exposure with academic achievement indicate that frequent readers are more successful students. Interestingly, poor readers also appear to benefit from independent leisure time reading. We conclude that shared book reading to pre-conventional readers may be part of a continuum of out-of-school reading experiences that facilitate children's language, reading, and spelling achievement throughout their development.

Based on:
Mol, S. E. & Bus, A. G. (in press). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin.*

Introduction

Popular media, governments, schools, and parents all encourage children to read in their leisure time. There is a widely held assumption that exposure to print makes us smarter and helps promote success in life. Is, however, this assumption supported by scientific evidence? Does reading for pleasure make us better and faster readers, more knowledgeable and even better speakers? How do the language and reading abilities of frequent readers differ from those of non-readers at each stage of development? To the best of our knowledge, there are no previous attempts that address these questions by synthesizing the evidence available across developmental levels.

Individual differences in print exposure are already present before any formal education, as parents vary in how often they read storybooks to their young children (Bus, 2001; Baker, Scher, & Mackler, 1997; Dickinson & McCabe, 2001; Heath, 1982; Mistry, Biesanz, Chien, Howes, & Benner, 2008; Raviv, Kessenich, & Morrison, 2004; Scheele, Leseman, & Mayo, 2010). We can regard parent-child book sharing as part of a continuum of leisure-time reading experiences that facilitate and influence reading skills throughout development. It seems plausible that variation in exposure to fiction books, magazines, comic books, and newspapers during leisure time increases with age. During the primary grades, children are mainly introduced to narrative texts, whereas their encounters with texts shift toward expository and technical texts from fourth grade onwards, as they must read to acquire knowledge in different content areas (RAND, 2002). Reading assignments for college and university students also include more non-fiction textbooks than narrative texts. Reading fiction books and the like, therefore, increasingly becomes a voluntary choice that entails additional and independent reading practice and, therefore, is likely to distinguish frequent and motivated readers from infrequent readers. Furthermore, because cognitive processing is enriched as a function of involvement, and because narratives are more likely than expository texts to stimulate imagination and to be personally relevant and/or emotionally engaging, the reading of fiction may especially support consolidation and extension of knowledge about word forms and word meanings (Hakemulder, 2000; Harding, 1962; Mar, 2004; Oatley, 1999). Reading narrative texts as a leisure-time activity may therefore have a different impact on reading skills across various ages and educational levels. This meta-analysis focuses on the role of print exposure during leisure time in reading development from infancy to early adulthood.

In essence, reading is the cognitive process of understanding speech that is written down. Young children form basic concepts about the connections between spoken and written words, leading to word recognition and familiarity with the spelling of words (Castles & Coltheart, 2004; Ziegler & Goswami, 2005). Initially, children develop alphabet knowledge (i.e., knowledge of letter names and how letters relate to sounds in spoken words), phonological processing skills (i.e., how words consist of separable sounds and the ability to manipulate phonemes), and orthographic processing skills (i.e., how to identify meaningful or frequently occurring parts in written words). These lower-order basic reading skills are considered to be the most time-constrained skills: After a period of rapid growth a ceiling is reached in the early primary grades (Paris, 2005; Paris & Luo, 2010). Likewise, technical reading and spelling skills may follow a similar time-constrained developmental trajectory, although it takes longer to reach mastery in word reading accuracy and fluency and in spelling words correctly. From early on, word reading ability may depend not only on basic reading skills but also on oral language skills such as vocabulary (e.g., Dickinson, McCabe, Anastasopolous, Peisner-Feinberg, & Poe, 2003; Oulette, 2006; Sénéchal & Cornell, 1993; Stanovich, 1986). As the ultimate goal of reading is reading for understanding, across development reading proficiency is less determined by technical reading skills and is more dependent on sophisticated vocabulary, background knowledge, and intelligence (e.g., Aarnoutse, Van Leeuwe, Voeten, & Oud, 2001; Hoover & Gough, 1990; Hulslander, Olson, Willcutt, & Wadsworth, 2010; Nation & Snowling, 2004; NRP, 2000; Storch & Whitehurst, 2002; Vellutino, Tunmer, Jaccard, & Chen, 2007).

In the current study, we address the claim that technical reading and spelling skills, as well as reading comprehension, are honed not only through direct instruction but also through print exposure. Furthermore, we examine whether leisure-time reading exerts an increasing impact on reading proficiency with growing age. The association between reading as leisure activity and the acquisition of reading skills may be an example of spiral causality or reciprocal causation (see Stanovich, 1986). When children enjoy reading books as a leisure-time activity, they read more often, which in turn improves both technical reading and spelling skills and reading comprehension, motivating children to continue reading (Cunningham, Stanovich, & West, 1994; Kush, Watkins, & Brookhart, 2005). As a result of increasing individual differences in leisure-time reading, we expect the relationship between print exposure and reading skills to strengthen across years of education.

Taking into account that technical reading and spelling skills have a relatively narrow window of learning and that only skills such as oral language and reading comprehension can be assessed at all ages (Paris & Luo, 2010), we conducted separate meta-analyses in three consecutive age groups: (a) preschoolers and kindergartners, (b) children in grades 1 to 12, and (c) undergraduate and graduate students attending college or university. We related print exposure to the following outcome domains: oral language (in particular expressive and receptive vocabulary), reading comprehension, and more general achievement measures as intelligence and academic achievement tests (e.g., eligibility test for university) as indicators of the *comprehension component;* and basic reading skills (alphabet knowledge, phonological processing, orthographic processing), word recognition (word identification, word attack), and spelling as indicators of the *technical reading and spelling component*.

Print Exposure and Comprehension

Book Sharing with Pre-Conventional Readers. Book reading is often seen as one of the most important activities for developing the knowledge required for eventual success in reading (Commission on reading, National Academy of Education, 1985; Samuelsson et al., 2005). Establishing a book-reading routine before the age of two is thought to provide children with a variety of rich linguistic input that stimulates their language development and lays the basis for continued, frequent print exposure (Duursma, 2007; Fletcher & Reese, 2005; Lyytinen, Laakso, & Poikkeus, 1998; Raikes et al., 2006). The metaphor of a "snowball" is used to illustrate how book sharing relates to language comprehension: As language develops due to book sharing, children's interest in books grows, thereby promoting linguistic exchanges with their caregivers that further refine word knowledge, syntax, and other aspects of language (Neuman, 2001; Raikes et al., 2006). Furthermore, starting to share books early is likely to optimize the quality of reading in the long term as frequent reading interactions may have the capacity to extend parents' knowledge of and sensitivity towards their children's linguistic and cognitive competencies (Fletcher & Reese, 2005). Such sensitive, high-quality interactions are likely to make reading more enjoyable for parent and child and lead to an increase in reading frequency, thereby increasing the likelihood for learning new language and expanding comprehension skills (Bus & Van IJzendoorn, 1988; De Jong & Leseman, 2001). In line with the "snowball" metaphor, we may expect a reciprocal effect in which comprehension skills develop as a result of exposure to books and in which comprehension determines whether children are exposed to book sharing.

Previous meta-analyses have supported the hypothesis that home literacy activities from an early age contribute substantially to young children's language and reading comprehension (Bus, Van IJzendoorn, & Pellegrini, 1995; Mol, Bus, De Jong, & Smeets, 2008; NELP, 2008). Children who have had storybooks read to them frequently – and who have parents who read themselves and own many books – enter school with larger vocabularies and more advanced comprehension skills than their peers who grow up in poorer home-literacy environments. A meta-analytic approach proceeds in a statistically rigorous way to analyze numerical results of studies with comparable outcome domains and variations in study characteristics (e.g., children's first language, mean age, socioeconomic status) (see Bus, Van IJzendoorn, & Mol, in press). Effect sizes, quantitative indexes of relations among variables, are used to compare and communicate the strength of the summarized research findings (Hedges, 2008). To ease interpretation, effect sizes can be converted into a Binominal Effect Size Display, which demonstrates the change in success ratio that can be attributed to the main variable of interest such as shared book reading (Rosenthal, 1991). For example, outcomes of the Bus et al.'s (1995) meta-analysis indicate that 64% of the children who are read to will be the more proficient readers at school compared to only 36% of children who are not exposed to books. This meta-analytic evidence is based not only on correlational studies but also on experimental and longitudinal research that allows for stronger causal inference. Therefore we could argue that book sharing makes a significant difference in children's lives by promoting knowledge and skills that are needed in order to learn how to read and by stimulating a positive attitude towards reading.

In a more recent set of studies than were included in Bus et al. (1995), the hypothesis was tested that book reading may in particular affect vocabulary acquisition, a central element of text comprehension (e.g., Dickinson & McCabe, 2001; Verhallen & Bus, 2010; Whitehurst & Lonigan, 1998). Children may learn more new words during reading than during other interactions with language, such as during mealtime and playtime, because children's books contain three times as many low-frequency words as do TV shows or adults' conversations with children (Hayes & Ahrens, 1988). Furthermore, caregivers may ask questions about pictures, difficult words, and story events, and give informative feedback on children's answers during book sharing, boosting story comprehension and language development (e.g., Collins, 2010; Mol et al., 2008; Mol, Bus, & De Jong, 2009; DeTemple & Snow, 2003; Whitehurst et al., 1988). Whether book reading results in receptive word learning (i.e., comprehending its meaning) as well as expressive word learning (i.e., producing the word) is still in debate. Some reading researchers show that expressive vocabulary may be promoted especially when children are challenged by caregivers to actively repeat or label words (Ard & Beverly, 2004; Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009; Penno, Wilkinson, & Moore, 2002; Sénéchal, 1997).

The present meta-analysis of print exposure in pre-conventional readers is an update as well as a critical replication. Research syntheses thus far may have systematically underestimated the effects of book sharing because studies assessed children's print exposure through self-report questionnaires. Parents are likely to overestimate the time they spend reading to their young children when they highly value book reading (DeBaryshe, 1995), which may reduce variance in questionnaire responses and attenuate the correlation between book reading frequency and comprehension measures. To test the impact of social desirability biases, we applied a cross-validation approach in order to directly compare (a) studies using traditional self-report questionnaires with (b) studies assessing parents' familiarity with children's book titles as measured by a print exposure checklist. The latter measure is more objective; it may reveal stronger correlations with language and story comprehension.

Independent Text Reading by Conventional Readers. Frequent exposure to texts broadens knowledge that enables readers to become more proficient in reading comprehension (e.g., Hirsch, 2003). In addition to general knowledge of the world, advanced levels of oral language skills are required for successful text comprehension. Independent text reading seems the most promising activity to develop such language skills; written texts not only contain a variety of words and complex sentence structures, but also provide context information that supports the readers' ability to infer meaning of unknown vocabulary (Nagy, 1988; Nagy & Hermann, 1987). However, readers need background knowledge as well as a mental lexicon that covers at least 95% of the words in a text to understand its content and to be able to guess unfamiliar words from context (Carver, 1994; Hu & Nation, 2000; Laufer, 1989). In line with a meta-analysis that showed that proficient readers and students in the upper grades have the greatest chance of incidental vocabulary acquisition (Swanborn & De Glopper, 1999), readers with smaller vocabularies are most likely to experience problems with understanding and learning vocabulary from age-appropriate texts.

When children lack background knowledge and vocabulary and therefore do not succeed in comprehending text, they become less eager to read, and, as a result, show stagnation in their reading comprehension skills, vocabulary size, and general knowledge base (Kush et al., 2005). Such a negative causal spiral could explain why reading development tapers off toward the end of fourth grade, when students are no longer learning to read by practicing relatively easy texts but must instead read to learn from subject-matter textbooks (Chall, 1983). Fourth-grade students are faced with texts that demand considerable oral language skills and efficient reading strategies to understand the content and to expand the knowledge base necessary to succeed in school (Hirsch, 2003; Juel, 2006; Vellutino et al., 2007). In contrast, an upward causal spiral may occur in proficient readers, who are more likely to have pleasurable reading experiences and who choose to read more often, resulting in continued improvements in language skills, background knowledge, and reading comprehension.

Differences in levels of print exposure may result in increasing inter-individual achievement differences over time for frequent readers versus infrequent readers, which is sometimes termed the "Matthew effect" (Bast & Reitsma, 1998; Foster & Miller, 2007; Stanovich, 1986). Such an achievement gap is likely to widen in particular for unconstrained skills such as oral language and reading comprehension, because learning new words and their meanings from context has few upper bounds. In other words, oral language and reading comprehension skills will continue to develop over the life span (Paris, 2005). Consequently, even among more proficient readers, individual differences in oral language skills, reading comprehension, and (possibly) intelligence and general academic achievement would be posited to increase as a function of print exposure (Stanovich, West, & Harrison, 1995; West, Stanovich, & Mitchell, 1993). We expect, therefore, that the correlations between print exposure and these unconstrained skills will get stronger as the number of years of education increases. Here too, we try to avoid the negative bias of self-report data by focusing on print exposure measures that are least sensitive to social desirability.

Print Exposure and Technical Reading and Spelling

Book Sharing and Basic Reading Skills. Children's storybooks may offer an incentive for the development of knowledge about print, letters, and sounds in pre-conventional readers, because storybook illustrations are mostly accompanied by the written text that parents can read aloud (Sulzby, 1985; Teale & Sulzby, 1986). Eye-tracking research shows that illustrations attract more visual attention than print (Evans & Saint-Aubin, 2005; Justice, Pullen, & Pence, 2008; Justice, Skibbe, Canning, & Lankford, 2005), but the proportion of time that children spend looking at the text during shared storybook reading increases from kindergarten to fourth grade and is greatest when the difficulty level of the text is within children's reading proficiency level (Roy-Charland, Saint-Aubin, & Evans, 2007). The youngest pre-conventional readers may pay barely any attention to print features in storybooks because they need all their working memory capacity to interpret the illustrations and to link the story content with the illustrations. Older children with more advanced basic knowledge about stories are more likely to notice and process print in storybooks even without their attention being drawn to print by their caregivers (De Jong & Bus, 2002; Evans, Saint-Aubin, & Landry, 2009; Neuman, 2001). We expect, therefore, a reciprocal relation between book sharing and basic reading skills, as storybooks promote the independent acquisition of print knowledge but only when some print knowledge is available.

Independent Text Reading and Technical Reading and Spelling. In narrative texts, words are presented in a relevant context, which may not only stimulate knowledge about the meaning of words but also improve word-reading skills in conventional readers (e.g., Krashen, 1989; Stanovich, 1986). Frequent encounters with words in context are assumed to strengthen basic reading skills and to lead to new connections between written word forms and syntactic and semantic information (Bowers, Davis, & Hanley, 2005; Ehri & Roberts, 1979; Pecher, Zeelenberg, & Wagenmakers, 2005). Apart from instructing and/or practicing single words, we suggest that text reading has at least two additional advantages. Reading words is not only more motivating when words are embedded in engaging stories (Guthrie & Wigfield, 1999), but the syntactical and semantic context can also be used to guess at less familiar words and to store, connect, and enrich associations between word forms and contextual information (Nation, 2008; Perfetti & Hart, 2002).

Basic reading skills. When children encounter unknown words while reading text, they follow the relatively slow *graphophonological route*. Beginning readers sound out individual letters and blend them into pronunciations that approximate real words (Ehri, 1998). They thereby improve lower-order reading skills via alphabet knowledge and phonological and orthographic processing of words. The self-teaching hypothesis predicts that applying letter-to-sound rules enables the acquisition of orthographic representations of novel words through independent print exposure (Jorm & Share, 1983; Share, 1995, 1999). As such basic reading skills typically evolve from nonexistent, to fully acquired, to automatic command in a restricted time span (Paris, 2005), we expect that the development of basic skills may benefit from print exposure especially in the primary grades. Poor readers seem to gain less word-specific knowledge from the same amount of print exposure than skilled readers (e.g., Breznitz, 1997; Ehri & Saltmarsh, 1995; Ehri & Wilce, 1979; Reitsma, 1983; Share & Shalev, 2004), and as a result, they take longer to master these constrained skills. Because poor readers will still vary in their basic reading skills while their peers with age-appropriate reading abilities are much more similar, the correlations between print exposure and basic reading skills are expected to be strongest for groups of poorer readers.

Word recognition. More advanced readers may increasingly process sound patterns of frequently occurring letter clusters and recognize the meaning of the blend (Ehri, 1998). In opaque languages such as English and French, applying letter-to-sound rules according to the *graphophonological route* is often not sufficient, because connections between letters or letter clusters and sounds are inconsistent (Goswami, Ziegler, Dalton, & Schneider, 2001; Patel, Snowling, & De Jong, 2004). Instead, advanced readers in such languages use the *lexicosemantic route,* where characteristics of the visual word form are directly associated with the word's meaning (e.g., Paulesu et al., 2000; Seymour, Aro, & Erskine, 2003). Low levels of print exposure are thought to delay the development of both the graphophonological and lexicosemantic routes that are required for adequate and fluent word recognition (Stanovich, Siegel, & Gottardo, 1997).

Reading words in context may be relevant especially for the development of orthographic representations of recurrent letter clusters (e.g., -ight), morphological patterns (e.g., -ed), or even higher order structures (e.g., whole words) that enable processing words through the lexicosemantic route (e.g., Ehri, 1998). Each exposure to a word embedded in a text sets down an "episodic trace" that relates word form information to the context in which the word occurred (e.g., pictures, events, sentences, other words). The episodic traces will be renewed each time the reader is confronted with the word form, further enhancing the quality of the lexical representation and contributing to the comprehension of the text that contains the word (see Nation, 2008; Shaywitz & Shaywitz, 2008). Because of an imbalance in print exposure levels among children, individual differences in the availability of episodic traces are likely to increase over time: Children who do not read much in their leisure time have lower quality representations of word forms and, hence, their development of word recognition is less advanced compared to frequent readers who repeatedly encounter word forms in a variety of contexts.

Spelling. The self-teaching hypothesis suggests that as a result of repeated encounters with words in written text, orthographic representations of word parts or complete words also contribute to writing skills (Cunningham, Perry, Stanovich, & Share, 2002; Share & Shalev, 2004). Children initially over-rely on phonetics when spelling dictated words, but as their development progresses they gradually move to strategies that incorporate sound, orthographic patterns, and semantics (Berninger et al., 2002; Bourassa & Treiman, 2001; Sadoski, Willson, Holcomb, & Boulware-Gooden, 2005). The complexity of English spelling and the lack of systematic teaching of morpheme-spelling rules in schools have led to the hypothesis that competent spellers infer spelling knowledge by reading, and not from training of spelling rules (Krashen, 1989; Nunes & Bryant, 2009). As even adults who are proficient in writing make spelling errors, we expect that spelling is less time-constrained than basic reading skills and word recognition, so its association with print exposure is likely to continue to become stronger with increasing years of education. For poor readers, however, it takes longer to acquire letter-to-sound rules which may interfere with learning word spellings, even when their amount of print exposure is comparable to that of more proficient readers (Ehri & Saltmarsh, 1995; Share & Shalev, 2004).

Reciprocal Causation?

Because of the correlational nature of the bulk of studies into print exposure, four possible interpretations of the association between reading abilities and print exposure may arise (e.g., Moore & McCabe, 2006). First, print exposure might be a causal factor in enhancing reading ability. For instance, book sharing is thought to support school readiness (e.g., Duursma, 2007; Wood, 2002) and the acquisition of conventional reading skills in the primary grades (e.g., McDonald-Connor, Son, Hindman, & Morrison, 2005; Melhuish et al., 2008; Molfese, Modglin, & Molfese, 2003). Second, print exposure may be largely a consequence of children's reading ability. Low-achieving readers may not perceive reading as a rewarding experience, which might result in less print exposure, whereas better readers are likely to have positive experiences with reading, which may be an incentive for reading as a leisure activity (e.g., Koolstra, Van der Voort, & Van der Kamp, 1997; Leppänen, Aunola, & Nurmi, 2005). Third, the association may be spurious due to lurking, or hidden, third variables, which are positively related to both reading skills and reading volume. A fourth possibility seems most plausible: Print exposure is both a consequence of reading ability and a contributor to further reading development, and the association may in fact be based on reciprocal causation (e.g., Bast & Reitsma, 1998; Harlaar, Dale, & Plomin, 2007). Overall, if print exposure makes a difference in children's (academic) lives, it may be expected that oral language skills, reading comprehension, basic reading skills, word recognition, spelling, and intelligence relate to the amount and frequency of reading for pleasure. Because more skilled readers are more likely to enjoy reading as a leisure-time activity, they will choose to read more frequently which, in turn, will improve knowledge of word forms and semantics, and enhance vocabulary size and text comprehension abilities.

As long as children are unable to read conventionally, they need caregivers who help them to bridge the gap between the world of the book and their own world (Bus, 2003). When children enter school and are no longer solely dependent on their caregivers for their print exposure, their home environment is still thought to explain achievement differences in the classroom (Alexander, Entwisle, & Olson, 2007; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996). However, the degree to which children evoke and select their own leisure time reading environment changes with development: As children mature, they may become more active creators of their own environments by seeking out stimulating experiences that are compatible with their abilities and interests. For children in preschool and kindergarten, their parents' behaviors will be the most critical element in determining their print exposure (e.g., Forget-Dubois et al., 2009), whereas for older children, their comprehension and technical reading and spelling skills will become more and more influential in whether they choose to read as a leisure activity, and the influence of their environment is likely to decrease (e.g., Harlaar et al., 2007; Petrill, Deater-Deckard, Schatschneider, & Davis, 2005). As children are not all equally attracted to reading fiction books, magazines, and the like, it seems probable that individual differences in leisure-time print exposure increase as children advance through the educational system.

Measurement of Print Exposure

The main inclusion criterion for the present meta-analysis was the administration of a print exposure checklist: an unobtrusive measure that is thought to be an objective proxy of reading volume (Stanovich & West, 1989; Stanovich, 2000). Print-exposure checklists follow a quick-probe logic in which titles of popular novels or names of best-selling authors function as probes into a person's literacy environment. The checklist can be adjusted to measure out-of-school reading in any age group by excluding titles or authors prominent in the school curriculum (e.g., Barker, Torgesen, & Wagner, 1992; Bråten, Lie, Andreassen, & Olaussen, 1999; Cunningham & Stanovich, 1997). Foils – fake items of non-existing titles or author names – are added to correct for guessing. It is assumed that a parent, child, or student who reads frequently will know more about literature and, therefore, will recognize more correct items than a respondent who reads less often (Allen, Cunningham, & Stanovich, 1992; Sénéchal, LeFevre, Hudson, & Lawson, 1996; West et al., 1993). Furthermore, the checklist is thought to reflect the attitude towards and familiarity with the domain of literature (Allen et al., 1994; Cunningham et al., 1994).

In previous qualitative (e.g., Evans & Shaw, 2008; Scarborough & Dobrich, 1994; Teale, 1981) and quantitative research syntheses (Bus et al., 1995), self-report questionnaires were included as the chief indicators of young children's exposure to print. Such questionnaires, however, are likely to suffer from a social desirability bias (DeBaryshe, 1995). In addition, many items are open to ambiguous interpretations and require retrospective time judgments (e.g., "How frequently have you read to your child in the past week?"). A parent might count the sharing of five books in one sitting before bedtime as five sessions, whereas another parent will report this as only one reading episode (Sénéchal et al., 1996). The literature even provides examples of parents who counted reading a word on a wrapper as a reading session (e.g., Van Lierop-Debrauer, 1990). Print-exposure checklists are thought to avoid these measurement issues and provide more objective insights in children's home literacy environment (Sénéchal et al., 1996).

We expect that the impact of measurement method will be greatest among parents of pre-conventional readers who may feel most inclined to overestimate their book reading frequency. With the media, pediatricians, and schools emphasizing that an early start with sharing storybooks ensures children's academic success, a questionnaire on book reading practices may feel like a "parental quality" test. Reporting that you do not manage to read daily is like admitting that you do not want to optimally prepare your child for school. In the set of studies on pre-conventional reading children, we therefore applied a cross-validation approach to test the impact of the expected bias. We compared two independent sets of studies that differed in the method they used to measure children's home literacy environment but that were comparable in their main study characteristics. That is, we matched each study in which parents completed a print-exposure checklist with a study that used a self-report questionnaire to assess young children's home literacy environment on characteristics such as sample size, children's mean age, home-language, and socioeconomic status. We expect that the self-report studies would replicate the main finding in earlier syntheses that about 8% of the variance in young children's language and reading comprehension is related to shared book reading (Bus et al., 1995; Scarborough & Dobrich, 1994). As print-exposure checklists are likely to be less biased, we expect that such checklists will reveal stronger correlations with outcome measures than will self-report questionnaires.

The Current Study

The meta-analysis presented here consisted of three steps. First, studies in which parents of preschoolers and/or kindergartners completed a print-exposure checklist were matched to studies that administered a self-report questionnaire. Second, we meta-analyzed studies linking print exposure to comprehension and technical reading and spelling skills of children attending grade 1 to 12. Third, as individual differences are predicted to increase until adulthood, we tested effect sizes for the relation between print exposure and all outcome domains within a set of studies on undergraduate and graduate students. In both groups of conventional readers (i.e., beyond preschool or kindergarten), we contrasted effects of print exposure in poorer readers against those found in their higher achieving peers. Specifically, we focused on the following hypotheses:

1) At all educational levels, indicators of the comprehension component (oral language, reading comprehension, or general achievement measures) as well as indicators of technical reading and spelling skills (basic reading skills, word recognition, or spelling) will be associated with print exposure.
2) For unconstrained skills such as oral language and reading comprehension, correlations with print exposure are expected to become stronger with increasing grade levels, because readers who have pleasurable reading experiences choose to read more often.
3) Constrained technical reading and spelling skills may remain correlated with print exposure for a longer period in low(er)-ability readers than in children with age-appropriate reading abilities.
4) For pre-conventional readers, effect sizes found in studies based on self-report questionnaires will be smaller than effect size estimates based on print-exposure checklists.

Method

Search Strategy and Inclusion Criteria

We entered into databases, such as PsycInfo, ERIC, and ProQuest Dissertations, several combinations of the following keywords: *print exposure, title/author/magazine recognition or checklist, home literacy environment, shared/joint/parent-child book reading, reading frequency, free voluntary reading, leisure time reading, reading development, reading ability, oral language, preschool, kindergarten, primary/elementary/middle/high school,* and/or *(college or university) students*. In addition, we read the method sections of articles that cited Stanovich and West (1989), Cunningham and Stanovich (1990; 1991), or Sénéchal et al. (1996) to check whether these citing studies used an (adapted) version of their print exposure checklists. We further extended our search by examining the reference lists of our included studies. As an additional check, we selected some representative journals (i.e., *Journal of Educational Psychology, Journal of Research in Reading, Reading Research Quarterly, Reading & Writing, Scientific Studies of Reading, Journal of Literacy Research,* and *Journal of Early Childhood Literacy Research*) and hand-searched journal issues from January 2004 to December 2008. We encountered no studies that we had not detected in our initial searches.

The selected articles had to meet the following inclusion criteria: (1) a print-exposure checklist had been administered, in which book titles, names of authors, and/or magazine titles were listed; (2) respondents were either parents of two- to six-year-old pre-conventional readers, school-aged children attending grade

1 to 12, or undergraduate and graduate students (studies assessing adults such as university staff were included only when the majority of the sample consisted of college or university students); (3) child outcome measures comprised oral language and/or reading ability tests and were administered in the same (school) year as the checklist(s) (studies that included only general measures such as a selection test for high school were excluded, as were studies that did not include an oral language measure in the group of pre-conventional readers); and (4) the correlations or means and standard deviations provided reflected the association between a print-exposure checklist and comprehension or technical reading and spelling outcomes and could be transformed into a Fisher's *z* effect size. There were no restrictions on study design or on participants' language or country, as long as the article did not report a case-study and was written in English, French, Dutch, or German. All (published or unpublished) articles, dissertations, or conference contributions were retrieved before January 2009.

We excluded print-exposure studies that reported no child outcomes or outcomes other than comprehension and technical reading and spelling skills, such as science tests or social ability tasks (e.g., Bråten et al., 1999; Burgess, 2005; Castles, Datta, Gayan, & Olson, 1999; Chomsky, 1972; Curry, Parrila, Stephenson, Kirby, & Catterson, 2004; Korat & Schiff, 2005; Lee & Krashen, 1996; Long & Prat, 2002; Mar, Oatley, Hirsh, Dela Paz, & Peterson, 2006; Pavonetti, Brimmer, & Cipielewski, 2003; Radloff, 2008; Stainthorp & Hughes, 1998), studies in which the checklist and the outcome measures were not administered within the same school year (e.g., Harlaar et al., 2007; Hood, Conlon, & Andrews, 2008; Shatil & Share, 2003; Stainthorp, 1997), and studies in which the participants were too old to meet our inclusion criteria (e.g., Lee, Krashen, & Tse, 1997; Stone, Fisher, & Eliot, 1999; West et al., 1993). Studies were also excluded when the respondents were teachers (e.g., McCutchen et al., 2002), kindergarten children (e.g., Bulat, 2005), or the parents of school-aged children (e.g., McGrath et al., 2007). Because mothers read most to the child, we utilized maternal data over paternal if both were reported (e.g., Symons, Szuskiewicz, & Bonnell, 1996). Attempts to locate the dissertation by Daly (2000), studying print exposure in 8-11 year-old children from Northern Ireland, were unsuccessful.

When multiple, independent samples were included within one article, we treated them as separate studies (Byrne, Fielding-Barnsley, Ashley, & Larsen, 1997; Ecalle & Magnan, 2008; Grant, Gottardo, & Geva, 2008; Grant, Wilson, & Gottardo, 2007; McBride-Chang, Manis, Seidenberg, Custodio, & Doi, 1993; Sears, Siakaluk, Chow, & Buchman, 2008; Stanovich & West, 1989) or we selected the sub-samples that met the inclusion criteria (Ecalle & Magnan, 2008; Sénéchal & LeFevre, 2002; Stanovich et al., 1995; Wolforth, 2000). The data from Burns and Blewitt (2000), Davidse, De Jong, Bus, Huijbregts, and Swaab (in press), Grant et al. (2008), Masterson and Hayes (2008), and Van der Kooy-Hofland, Kegel, and Bus (in press) were obtained by e-mailing the author(s).

To cross-validate the print-exposure checklist in the group of pre-conventional readers, we matched the studies in which parents filled in a print-exposure checklist with studies that administered only a self-report questionnaire about parents' literacy resources and/or activities. Because correlations are influenced by sample size (Lipsey & Wilson, 2001; Moore & McCabe, 2006), we searched databases and abstracts for studies with comparable samples. For each print-exposure study included, we then tried to find a match on four main characteristics: sample size, children's mean age, home language, and socioeconomic status. Except for one study with 24 English-speaking preschool children from India (Kalia, 2007), we were able to match each of the 15 studies with a comparable counterpart (see Appendix 2.1 and 2.2). This cross-validation approach gave us the unique opportunity to independently study differential effects of two measurement methods.

Coding Process

Two independent coders completed a standard coding scheme per study, comprising (a) year of publication; (b) publication status (published in peer-reviewed journal, unpublished, dissertation); (c) continent (Asia, Australia, Europe, North America) and specific country; (d) design (cross-sectional and/or longitudinal, (quasi-)experiment), (e) sample size and number of boys/girls; (f) mean age and age range; (g) socioeconomic status (low, middle-high); (h) school type (preschool, kindergarten, elementary/middle/high school (specify grade number), undergraduate, graduate, combination); (i) ability level (low(er) ability, age-appropriate, high(er) ability); (j) language learners (first, second); (k) print exposure checklist characteristics (language, number of (real and fake) items, composition procedure, scoring, Cronbach's α); (l) home literacy questionnaire (administered: yes, no; content of questions); (m) type and names of outcome measure(s) (standardized, unstandardized); and (n) correlation (bivariate, partial). Two coders coded seventy-five percent of all studies included. The intercoder agreement for both study characteristics and outcome variables ranged between 77% and 100% across meta-analyses, resulting in an overall average of M = 94.5% (κ = .96, *range* = .65 – 1.00). All discrepancies between coders were settled in discussion and consensus scores were used.

Because it can be assumed that standardized measures are more reliable and valid than unstandardized measures, we first treated standardized and unstandardized measures separately to check for differences in correlations with print exposure. Unconstrained skills such as *Oral Language* were assessed by standardized measures such as the PPVT or vocabulary subtests from the Metropolitan Achievement Test and the Nelson-Denny Reading Test. Vocabulary checklists (i.e., ticking off actual words in a list that also includes non-existent words) were treated as unstandardized. *Reading Comprehension* was predominantly measured by standardized tests that had children read short passages and answer multiple-choice or open-ended questions or fill in missing words in a cloze task: the Stanford Diagnostic Reading Test, Iowa Tests of Basic Skills, Neale Analysis of Reading Ability, Nelson-Denny Reading Test, Woodcock-Johnson Passage Comprehension, Peabody Individual Achievement Test, or the Stanford Early School Achievement Test. Constrained skills such as *Alphabet Knowledge* (e.g., naming letters), *Phonological Processing* (e.g., choosing one out of two pseudo-words that can be pronounced as a real word), and *Orthographic Processing* (e.g., pick the correct spelling from two choices that sound alike) were mostly measured by unstandardized tests and were treated as components of *Basic Reading Skills*. *Word Recognition* tests were separately coded as *Word Identification* (e.g., the ability to correctly identify words in isolation) and *Word Attack* (e.g., reading aloud pseudo-words and/or exception words), which were measured by standardized tests as the Woodcock-Johnson, Woodcock Reading Mastery Test, or the Test of Word Reading Efficiency. *Spelling* was assessed by standardized tests as the Wide Range Achievement Test, or by unstandardized experimental tasks such as writing dictated words. Error rates were preferred; reading speed measures or decision latencies were excluded. We also coded measures of IQ (i.e., RAVEN, WISC, Stanford-Binet) and indicators for academic achievement as the Grade Point Average (GPA), American College Testing (ACT), and Scholastic Assessment Test (SAT) scores.

Meta-Analytic Procedures

All correlations between a print exposure checklist and any outcome variable were inserted into the computer program Comprehensive Meta-Analysis (Borenstein, Hedges, Higgins, & Rothstein, 2005) and transformed into Fisher's z effect sizes for further analyses, because the variance of z' is approximately constant whereas the variance of the correlation follows an asymmetrical distribution (Borenstein, Hedges, Higgins, & Rothstein, 2009). To ease interpretation of the result section, Fisher's z summary estimates were transformed back into a correlation r with the formula: $r = \tanh(z')$ (Lipsey & Wilson, 2001). In general, a Fisher's z value of $z' = .10$ ($r = .10$) can be interpreted as a small effect size, $z' = .31$ ($r = .30$) as moderate, and $z' = .55$ ($r = .50$) as a large effect size (Cohen, 1988).

For studies that did not report bivariate Pearson r's we converted the provided statistics into Fisher's z values. A p-value of $p = .10$ was entered and converted into a weighted correlation for studies that only reported that an association was not significant. Kalia (2007), however, reported the range of non-significant correlations, so we entered $p = .50$ for all non-significant values to estimate a conservative correlation in the lower end of that range. Studies in which partial correlations ($k = 11$), converted F- and t-tests ($k = 4$), or means and standard deviations ($k = 8$) were provided were scattered through all outcome measures and did not influence the results when we analyzed the data without them.

To compare the effect sizes of print exposure for different outcome domains (oral language, reading comprehension, general achievement, basic reading skills, word recognition, spelling), we treated each outcome domain as an independent correlate (see Bus et al., in press). When a study utilized multiple tests to measure one outcome domain, we averaged the effect sizes within that study to ensure that each study contributed only one effect size to the analysis of that domain so that each had an equal impact on the summary estimate of each domain. For oral language, reading comprehension, and spelling skills, our stepwise approach included: (1) aggregating effects of standardized and unstandardized tests into two separate composites; and (2) if both were available, combining the standardized and unstandardized composites to create an overall composite per study. As basic reading skills were mostly measured by unstandardized tests and word recognition and general achievement by standardized tests, we did not distinguish standardized from unstandardized composites in these analyses. For each study that assessed more than one indicator of lower-order technical reading skills, we (1) created separate composites of alphabet knowledge, phonological processing, and orthographic processing per study; and (2) integrated these indicators into a basic reading skills composite. Likewise, combined effects for word identification and word attack were first calculated and then aggregated into a word-recognition composite that reflects higher-order or conventional technical reading skills. As far as the articles had not presented a composite for the print exposure checklists, we merged the title- and author-recognition test per outcome domain within the sample of preschool and kindergarten children, and the title-, author-, and magazine-recognition tests for the children in grade 1 to 12.

Samples were coded as "low(er) ability" when it was explicitly stated that students were reading disabled, had special-educational needs, or were in the lower third of a distribution that was based on a large set of students. Studies comprising second-language learners who were not tested in their first language were also treated as "low(er) ability". When groups of students were matched on a reading ability measure, the skill on which the groups were selected to differ was treated as the outcome variable. For example, Ricketts, Nation, and Bishop (2007) matched 15 poor and 15 skilled reading comprehenders on age, nonverbal ability, and decoding level, and administered an author recognition test. We transformed the checklist means and standard deviations of both groups into a Fisher's z and treated reading comprehension as the outcome variable, because the groups had been selected to differ significantly on reading comprehension. Because we analyzed both word recognition and reading comprehension as separate outcome variables, we had to exclude one subgroup in Leach, Scarborough, and Rescorla (2003) that showed combined deficits in word-level and reading comprehension skills. For all moderators and aggregated outcomes per study, see Appendix 2.3 and 2.4.

To estimate the mean effect size, we applied the conservative random-effects model in which studies are weighted by the inverse of their variance and, in addition, within-study error and between-study variation in true effects are accounted for (Borenstein et al., 2009). A combined effect, the precision of which is addressed by the 95% confidence interval (CI), is considered significant if the CI does not include zero. Differences between estimates are interpreted as significant when the CIs do not overlap. To avoid lack of power in the detection of meaningful differences across subgroups (Hedges & Pigott, 2004), a significant $Q_{\text{between}}(df)$ value for moderator analyses was only interpreted if the smallest subgroup contained a minimum of four studies (see Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003; Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van IJzendoorn, 2007).

Because studies with significant findings are more likely to be published and, therefore, are more likely to be included in a meta-analysis than unpublished studies, we examined whether the results were moderated by publication status. To the extent that the subgroups could be contrasted, published studies did not reveal significantly different correlations than unpublished studies (pre-conventional readers (matched set): $Q(1)_{HLE\text{-}comp*Basics} = 3.27$, $p > .05$; college and university students: $Q(1)_{ART*Oral} = 1.42$, $p > .05$, $Q(1)_{MRT*Oral} = 1.71$, $p > .05$, $Q(1)_{ART*WordRec} = 1.23$, $p > .05$). As another indicator, we calculated Rosenthal's fail-safe number (*N*fs), which reflects the number of missing studies with null effects that would have to be retrieved and included in the analyses before the *p*-value becomes non-significant (Borenstein et al., 2009). Because effects can be negligible but still significant, we also inspected funnel plots in order to address the potential impact of a publication bias. We reported adjusted effect sizes based on the trim-and-fill approach if there appeared to be asymmetry around the point estimate (Duval & Tweedie, 2000a, 2000b). In the current meta-analyses, 23 out of 79 summary point estimates had to be adjusted slightly, with a maximum of 3 imputed studies to the left of the mean ($m_{adjustment\ z'} = -.03$, *range* = -.01 – -.09). Overall, standardized *z* values fell within the range of -3.26 to 3.26 for all effect sizes ($p < .001$), implying that no outliers were present.

Results

The results of the meta-analyses are presented in six sections. First, we report study and sample characteristics. Second, we explore interrelations between measurement methods of print exposure in all age groups. In other words, we examine whether print exposure checklists correlated with scores on self-report questionnaires that contained items such as reading frequency, the number of books at home, and/or activity preferences (e.g., "I would rather read than listen to music of my choice"). In three subsequent subsections, we present correlations between print exposure and comprehension and technical reading and spelling outcomes for (a) preschool and kindergarten children, (b) children attending grade 1 to 12, and (c) undergraduate and graduate students. Across these three subsections, the effect sizes of oral language and reading comprehension are reported first, followed by the effect sizes of technical reading and spelling skills such as basic reading skills, word recognition, and spelling. In addition, results of meta-regressions and moderator analyses are presented. In the sixth and final section, longitudinal studies are reviewed to examine the plausibility of reciprocal causation.

For reasons of clarity, we report which mean effect sizes differed significantly from other mean effect sizes (i.e., the 95% CIs do not overlap) without mentioning the specific CIs in the text. These details as well as weighted combined effect sizes for the separate outcome variables of each domain can be found in Tables 2.1-2.4.

Descriptive Statistics

Ninety-nine studies (N = 7,669) met our inclusion criteria, of which 81 were published in peer-reviewed journals. Specifically, 29 studies comprised preschool and kindergarten children (n = 2,168), 40 studies targeted children attending grades 1 through 12 (n = 2,792), and 30 studies included undergraduate and graduate students (n = 2,709). Most respondents resided in North America ($k_{P\&K}$ = 24, n = 1,837; $k_{Gr1\text{-}12}$ = 27, n = 1,889; $k_{Students}$ = 24, n = 2, 219), were first language learners ($k_{P\&K}$ = 26, n = 1,777; $k_{Gr1\text{-}12}$ = 33, n = 2,368; $k_{Students}$ = 30, n = 2,709), and were tested in English ($k_{P\&K}$ = 21, n = 1,448; $k_{Gr1\text{-}12}$ = 36, n = 2,515; $k_{Students}$ = 29, n = 2,690). Information on socioeconomic status or parental education levels was only available for the youngest group of pre-conventional readers: Thirteen out of 15 homes in which the print exposure checklists were administered, and 11 out of the 14 matched studies, could be classified as middle-to-high socioeconomic status.

Correlations of Print Exposure Checklists and Home Literacy Questionnaires

Parents of preschoolers and kindergartners completed a child-title recognition test to assess familiarity with titles of children's storybooks (k = 13, n = 980), a child-author recognition test that lists authors of children's storybooks (k = 7, n = 576), and/or an adult-author recognition test comprising authors of adult fiction (k = 8, n = 658). Children in grade 1 to 12 mostly completed a title recognition test (k_{TRT} = 32, n = 2,311; k_{ART} = 14, n = 1,087; k_{MRT} = 7, n = 394), whereas undergraduate and graduate students all completed an author recognition test (k_{TRT} = 1, n = 80; k_{ART} = 30, n = 2,709; k_{MRT} = 17, n = 1,630). Overall, print exposure checklists contained more true items than foils ($m_{total\ items}$ = 51.94, sd = 29.78, $range$ = 8 – 150; $m_{\%true\ items}$ = 60.65%, sd = 10.35), and showed good mean reliabilities ($range\ m_{Cronbach's\ \alpha}$ = .75 – .89). As can be seen in Table 2.1, parents' knowledge of adult fiction correlated rather strongly with their knowledge of children's literature (r = .48, p < .001). Within the set of students, the author recognition test correlated strongly with the magazine recognition test (r = .60, p < .001).

Table 2.1
Interrelations between Print Exposure Checklists and Home Literacy Questionnaires across Meta-Analyses.

	Children's Literature (CAR+CTR)						Adult Fiction (AAR)					
Preschool/Kindergarten	*k*	Fisher's *z*	95% CI	*Q*	I^2	*Nfs*	*k*	Fisher's *z*	95% CI	*Q*	I^2	*Nfs*
Adult Fiction (AAR)	4	.52***	.32, .72	14.06**	78.66	13						
Freq. Reading to Child	8	.22***	.14, .30	8.90	21.36	66	4	.14	-.00, .28	4.50	33.27	2
Number of Books at Home	5	.50***	.42, .58	1.46	.00	172	4	.36***	.25, .47	0.25	.00	35

	ART					
Grade 1-12	*k*	Fisher's *z*	95% CI	*Q*	I^2	*Nfs*
Magazine RT	3					
HLE-Composite	5	.23**	.06, .39	16.49**	75.74	34

	ART						MRT					
(Under)Graduate Students	*k*	Fisher's *z*	95% CI	*Q*	I^2	*Nfs*	*k*	Fisher's *z*	95% CI	*Q*	I^2	*Nfs*
Magazine RT	14	.70***	.59, .81	36.79***	67.39	1,662						
HLE-Composite	6	.40***	.34, .47	4.54	.00	178	5	.25***	.16, .34	9.40	.83	66
Activity Preference = Reading	5	.48***	.38, .57	4.93	18.82	139	4	.24***	.15, .34	1.12	.00	16
Activity Preference = TV	4	-.18*	-.34, -.02	5.41	44.59	9	3					

Note. k = number of studies; *95%* CI = Confidence Interval; non-significant Qs imply homogeneity ($df = k$-1); I^2 reflects the degree of inconsistency among studies; *Nfs* = failsafe number; *HLE* = Home Literacy Environment; *CAR+CTR* = Child-Author Recognition and Child-Title Recognition Test; *AAR* = Adult-Author Recognition Test; *ART* = Author Recognition Test, *MRT* = Magazine Recognition Test; *** $p < .001$, ** $p < .01$, * $p < .05$

A small subset of studies also administered a self-report home literacy environment questionnaire ($k_{P\&K}$ = 10, n = 783; $k_{Gr1\text{-}12}$ = 5, n = 445; $k_{Students}$ = 8, n = 770) and/or an activity preference questionnaire with forced-choice questions that contrasted reading as well as television with other leisure time activities ($k_{P\&K}$ = 0; $k_{Gr1\text{-}12}$ = 2, n = 90; $k_{Students}$ = 5, n = 634). With parents as respondents, the number of books at home was significantly more strongly related to knowledge of children's literature (r = .46, p < .001) than a single item about the frequency of shared book reading (r = .22, p < .001) as appeared from non-overlapping 95% CIs. The correlations between undergraduate and graduate students' print-exposure checklist scores and activity-preference scores for reading were significantly higher for the author recognition test (r = .45, p < .001) than for the magazine recognition test (r = .24, p < .001). In the same vein, the author recognition test (r = .38, p < .001) was more strongly related to the home literacy composite than the magazine recognition test (r = .25, p < .001). Interestingly, a preference for television viewing correlated negatively with a students' score on the author recognition test (r = -.18, p < .05).

Meta-Analysis 1: Preschool and Kindergarten Children

In the set of two- to six-year-old children (M_{age} = 56.95 months, SD = 10.40), the correlation between oral language skills and print exposure checklists of children's literature was moderate (k = 12, r = .34, p < .001). An additional 478 non-significant studies would be needed to transform this significant result into a non-significant effect size (see Table 2.2, which presents fail-safe numbers for the effect sizes presented hereafter). Similar, moderate correlations were found for receptive (k = 9, r = .33, p < .001) and expressive vocabulary skills (k = 4, r = .35, p < .001).

To compare these effect sizes with a matched set of studies in which only a home literacy self-report questionnaire was administered, we calculated the weighted average with a composite of home literacy questions and the frequency of shared book reading as a single item in 14 studies that resembled the print exposure studies in terms of number of children, mean age, home language, and socioeconomic status. First, the correlations between oral language and the home literacy composite in matched studies (k = 11, r = .32, p < .001) were significantly stronger than the correlations with the frequency of shared book reading in matched studies (k = 6, r = .16, p < .01). Within the set of print-exposure studies, the same pattern was present when comparing the effect sizes for print-exposure checklists on children's literature with a single question about parent-child reading frequency (k = 8, r = .21, p < .001), whereas parents' estimation of the total number of books at home (k = 5, r = .32, p < .001) revealed almost identical correlations with oral language as print exposure checklists. Second, when we contrasted the matched self-report studies with the set of print exposure studies, the home literacy composite revealed similar combined effect sizes with oral language to the

Table 2.2

Effect Sizes between Print Exposure and Language and Basic Reading Outcomes for the Checklist-Studies and the Matched Self-Report Questionnaire Studies in Preschool and Kindergarten.

Print Exposure Studies		*Oral*	RV	EV	*Basics*	AK	PP	OP
Checklist								
Children's Literature (CAR+CTR)	*k*	12	9	4	8	5	8	2
	z'	.35***	.34***	.36***	.30***	.26***	.28***	
	95% CI	.27, .42	.26, .43	.22, .51	.22, .38	.18, .36	.21, .36	
	Q	19.13	11.84	5.29	13.29	2.80	6.49	
	I^2	42.48	32.41	37.23	47.31	.00	.00	
	*N*fs	478	224	29	222	35	102	
Adult Fiction (AAR)	*k*	8	6	3	5	1	4	4
	z'	.27***	.29***		.27***		.27***	.20
	95% CI	.20, .33	.19, .39		.21,.34		.17, .36	-.01, .40
	Q	7.20	8.14		2.77		.40	10.62*
	I^2	2.72	26.5		.00		.00	71.74
	*N*fs	123	62		73		25	12
HLE Questionnaire								
item: Frequency Reading to Child	*k*	8	7	2	4	2	3	1
	z'	.21***	.19***		.28***			
	95% CI	.13, .29	.11, .28		.18, .39			
	Q	7.72	3.45		2.66			
	I^2	9.28	.00		.00			
	*N*fs	60	25		22			
item: Number of Books at Home	*k*	5	4	2	2	1	2	1
	z'	.33***	.34***					
	95% CI	.24, .43	.22, .46					
	Q	3.72	3.58					
	I^2	.00	16.22					
	*N*fs	52	35					
Matched Studies								
HLE Questionnaire								
Composite-Scale	*k*	11	8	6	13	10	6	0
	z'	.33***	.35***	.33***	.18***	.19***	.21***	
	95% CI	.27, .40	.22, .48	.22, .43	.12, .24	.10, .28	.15, .27	
	Q	12.94	15.64*	3.29	29.08*	28.85*	4.44	
	I^2	22.69	55.24	.00	34.88	48.30	.00	
	*N*fs	372	119	60	287	162	49	
item: Frequency Reading to Child	*k*	6	5	3	7	3	4	0
	z'	.16**	.15**		.18***		.17**	
	95% CI	.10, .22	.06, .24		.11, .24		.07, .26	
	Q	.68	.94		2.33		.10	
	I^2	.00	.00		.00		.00	
	*N*fs	28	9		37		7	

Note. Oral = Oral Language Composite, *RV* = Receptive Vocabulary, *EV* = Expressive Vocabulary, *Basics* = Basic Reading Composite, *AK* = Alphabet Knowledge, *PP* = Phonological Processing, *OP* = Orthographic Processing; *HLE* = Home Literacy Environment; *CAR*+CTR = Child-Author and Title Recognition Checklist; *AAR* = Adult-Author Recognition Checklist; *k* = number of studies; *95%* CI = Confidence Interval; non-significant *Q*s imply homogeneity ($df = k$-1); I^2 reflects the degree of inconsistency among studies; *N*fs = failsafe number; *** $p < .001$, ** $p < .01$, * $p < .05$

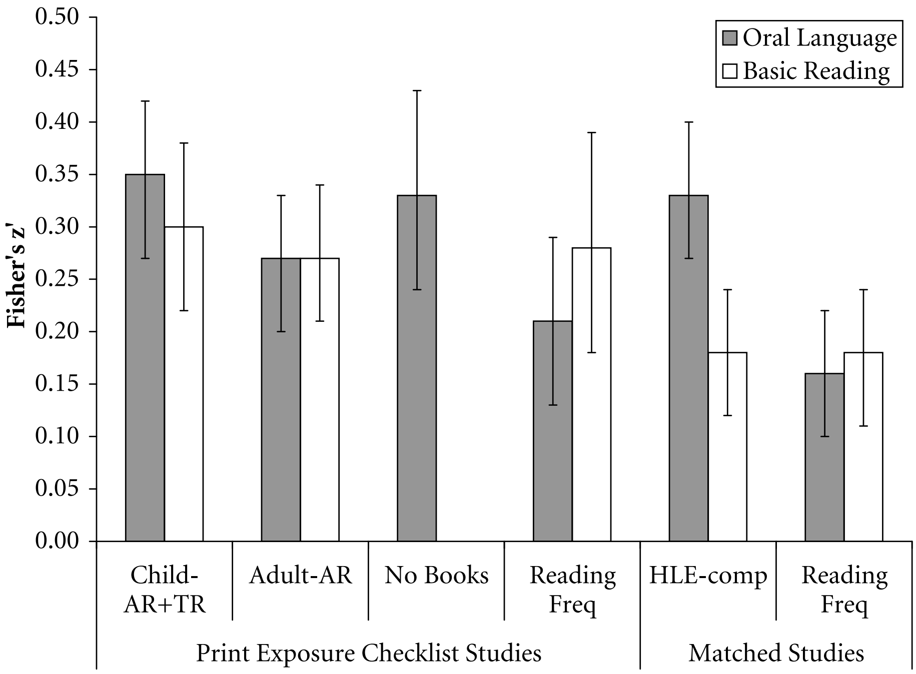
set of print-exposure studies. In sum, both composite scores of children's home literacy environment and print-exposure checklists are related moderately strong to oral language.

Print exposure showed a moderate effect size for basic reading skills as well (k = 8, r = .29, p < .001) and the 95% CI showed overlap with the CI of oral language. The set of matched studies revealed small correlations with the basic reading composite ($k_{HLE\text{-}Comp}$ = 13, r = .18, p < .001; k_{rfreq} = 7, r = .18, p < .001), and these were significantly smaller than for oral language, given non-overlapping CIs (see Figure 2.1).

Unfortunately, it was not possible to study age effects by contrasting preschool and kindergarten children or entering M_{age} into a meta-regression, because 7 studies included large, overlapping age ranges. Outcomes of print exposure studies that were carried out by Sénéchal and colleagues, who carried out nearly half of all studies with the checklist for children's literature (k = 5), did not significantly differ from studies from other research groups ($Q_{Oral}(1)$ = .20, p > .05).

Figure 2.1

Print Exposure Checklist versus Matched Set of Studies and their Effect Sizes for various Home Literacy-Indicators with Oral Language and Basic Reading Skills in Preschool and Kindergarten.



Note. Child-AR+TR = Child-Author and Title Recognition Checklist; *Adult-AR* = Adult-Author Recognition Checklist; *No Books* = Number of Books at Home (single item); *Reading Freq* = Reading Frequency (single item); *HLE-comp* = Composite of Home Literacy Environment questionnaire

Meta-Analysis 2: Grade 1 to 12

For children between 6.2 and 17.5 years of age (M_{age} = 10.23, *SD* = 2.61), the effect sizes between print exposure and all outcome measures ranged between r = .15 and r = .45. Standardized and unstandardized tests revealed comparable results and are presented as a composite here (see Table 2.3 for separate estimates).

Overall, print exposure was moderately related to oral language skills (k = 18, r = .45, p < .001) and to reading comprehension (k = 21, r = .36, p < .001). Second, moderate effect sizes for word recognition (k = 24, r = .38, p < .001) and spelling (k = 9, r = .42, p < .001) differed significantly from the smaller summary estimates that were found for basic reading skills (k = 18, r = .23, p < .001). The 95% CIs for oral language skills, word recognition, and spelling did overlap, whereas oral language did significantly differ from basic reading skills. In addition, IQ (k = 8, r = .15, p < .05) seemed to be affected significantly less by print exposure than oral language, reading comprehension, word recognition, and spelling.

In order to test whether the effect sizes between print exposure and outcome measures would be higher as a function of age, we conducted meta-regression analyses by entering M_{age} as a continuous variable. The random model (method-of-moment) meta-regression was significant for oral language (Q_{model} = 5.31, p < .05, $B_{(slope)}$ = .04), basic reading skills (Q_{model} = 7.63, p < .01, $B_{(slope)}$ = .03), and IQ (Q_{model} = 9.48, p < .01, $B_{(slope)}$ = .06), implying (if longitudinal reasoning could be applied to these cross-sectional data) that children gain z' = .04, z' = .03, and z' = .06 points each year as they get older, respectively, which will result in an increase of .36 to .72 standard deviations in the course of 12 years. Furthermore, the slopes of reading comprehension (Q_{model} = 2.92, p = .09, $B_{(slope)}$ = .04) and spelling skills (Q_{model} = 3.22, p = .07, $B_{(slope)}$ = .04) approached significance, whereas there was no such a trend for word-recognition (Q_{model} = .09, p > .50). Because a small number of studies might bias the results of regressions (Borenstein et al., 2009), we also conducted moderator analyses in which we categorized children's grades into primary (grade 1-4), middle (grade 5-8), and high school (grade 9-12). It should be noted that studies assessing high school students could only be included in the analysis for oral language, as the other skills were not typically assessed for them. Significant grade differences were present for oral language ($Q(2)$ = 11.81, p < .01; $k_{primary}$ = 6, r = .36, p < .001; k_{middle} = 7, r = .44, p < .001; k_{high} = 4, r = .55, p < .001) and word recognition ($Q(1)$ = 4.34, p < .05; $k_{primary}$ = 16, r = .31, p < .001; k_{middle} = 5, r = .48, p < .001), but did not appear for basic reading skills ($Q(1)$ = 2.18, p > .05) and reading comprehension ($Q(1)$ = 2.29, p > .05). In short, the correlations between print exposure and oral language were progressively stronger at higher levels of education. This pattern also seemed to emerge for technical reading skills and IQ from primary to middle school.

We also contrasted studies that contained children with age-appropriate abilities with studies that tested children with low(er) reading abilities. In line with our third hypothesis, no ability-level differences were detected for unconstrained skills such as oral language ($Q(1) = 1.14$, $p > .05$) and reading comprehension ($Q(1) = .01$, $p > .05$). However, the correlations between print exposure and basic reading skills were significantly stronger ($Q(1) = 9.57$, $p < .01$) for children with low(er)-ability levels ($k = 7$, $r = .39$, $p < .001$) than for children with age-appropriate reading abilities ($k = 11$, $r = .20$, $p < .001$), a distinction that was not detected for word recognition ($Q(1) = .57$, $p > .05$).

Table 2.3
Effect Sizes for the Print Exposure Checklists (Author, Title, and Magazine Recognition Tests) and All Outcome Measures for Meta-Analysis 2: Grade 1 to 12.

	k	Fisher's *z*	*95%* CI	*Q*	I^2	*N*fs
Oral Language	18	.49***	.42, .56	25.13	32.34	1,339
Standardized Tests	11	.43***	.36, .50	8.94	.00	332
Unstandardized Tests	11	.55***	.44, .66	18.59*	51.59	535
Reading Comprehension	21	.38***	.27, .50	88.35***	77.36	994
Basic Reading Skills	18	.23***	.16, .29	31.82	30.95	341
Alphabet Knowledge	2					
Phonological Processing	14	.22***	.14, .29	18.98	31.52	152
Orthographic Processing	6	.34***	.21, .46	4.74	.00	52
Word Recognition	24	.40***	.30, .50	122.79***	81.27	1,936
Word Identification	22	.42***	.32, .53	99.91***	77.98	1,815
Word Attack	9	.22***	.11, .33	15.33	34.24	68
Spelling	9	.45***	.32, .58	32.97***	75.73	459
Standardized Tests	3					
Unstandardized Tests	7	.48***	.37, .59	10.78	44.34	261
General Achievement						
IQ	8	.15*	.03, .26	15.47	44.82	26

Note. k = number of studies; *95%* CI = Confidence Interval; non-significant *Q*s imply homogeneity ($df = k$-1); I^2 reflects the degree of inconsistency among studies; *N*fs = failsafe number; *** $p < .001$, ** $p < .01$, * $p < .05$

Meta-Analysis 3: Undergraduate and Graduate Students

In the set of 30 studies comprising college and university students (M_{age} = 21.00 years, $SD = 2.32$), 17 included both author- and magazine recognition tests to measure print exposure. Overall, author recognition tests showed stronger correlations with all outcome variables than the magazine recognition tests: 95% CIs did not overlap for spelling outcomes and hardly showed any overlap for the other skills (see Table 2.4). In this section, therefore, we focus on author recognition checklists as the indicator of print exposure. We did not detect any significant differences between standardized and unstandardized tests, so we present only composites.

Table 2.4

Effect Sizes for Author and Magazine Checklists and all Outcome Measures for Meta-Analysis 3: Undergraduate and Graduate Students.

	Author Recognition Test						*Magazine Recognition Test*					
	k	Fisher's *z*	95% CI	*Q*	I^2	*Nfs*	*k*	Fisher's *z*	95% CI	*Q*	I^2	*Nfs*
Oral Language	18	.66***	.57, .74	38.25**	55.55	2,581	11	.46***	.32, .59	27.36**	63.45	409
Standardized Tests	12	.67***	.58, .75	15.08	27.07	1,102	8	.48***	.32, .63	14.01	50.05	225
Unstandardized Tests	7	.56***	.37, .76	46.85***	81.04	455	4	.43***	.19, .66	12.58**	76.15	53
Reading Comprehension	11	.44***	.33, .55	31.93***	68.68	644	6	.41***	.30, .53	10.43	52.04	168
Basic Reading Skills	6	.24***	.15, .33	12.32*	59.43	105	2					
Phonological Processing	5	.19***	.11, .27	7.11	25.65	38	2					
Orthographic Processing	5	.26***	.15, .37	8.16	50.98	54	1					
Word Recognition	9	.35***	.26, .43	9.89	19.10	178	5	.20***	.13, .28	.87	.00	21
Word Identification	8	.39***	.28, .51	12.13	42.27	158	4	.24***	.12, .35	.55	.00	9
Word Attack	6	.36***	.29, .44	5.34	6.43	107	2					
Spelling	14	.42***	.33, .51	26.75*	51.40	651	8	.20**	.08, .31	12.76	45.15	38
Standardized Tests	9	.43***	.33, .54	14.26	43.89	265	7	.25***	.15, .35	7.80	23.07	42
Unstandardized Tests	9	.37***	.27, .46	12.67	36.86	220	4	.06	-.03, .15	3.85	.00	1
General Achievement												
IQ	6	.18*	.00, .35	25.01**	73.92	43	5	.34***	.26, .42	3.67	.00	46
Academic Achievement	10	.31***	.21, .41	21.56*	62.90	211	8	.28***	.14, .42	26.17**	73.28	111

Note. k = number of studies; *95%* CI = Confidence Interval; non-significant *Q*s imply homogeneity (*df* = *k*-1); I^2 reflects the degree of inconsistency among studies; *N*fs = failsafe number; *** $p < .001$, ** $p < .01$, * $p < .05$

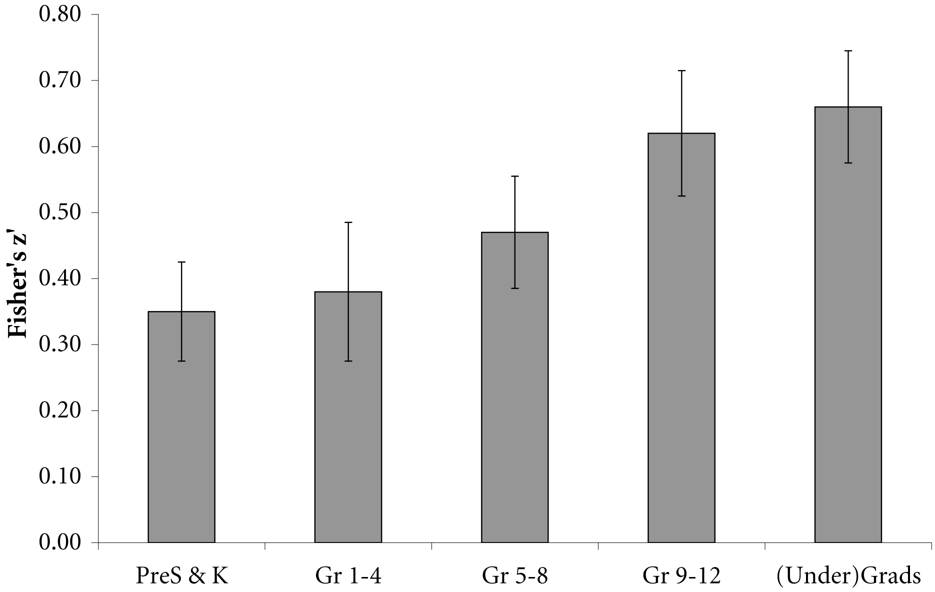
Oral language skills showed strong correlations with print exposure (k = 18, r = .58, p < .001), yielding a significantly stronger association than the moderate effect size found for reading comprehension (k = 11, r = .41, p < .001) as no overlap was detected between 95% CIs. Technical reading and spelling skills were small to moderately related to print exposure (k_{Basics} = 6, r = .24, p < .001; $k_{WordRec}$ = 9, r = .34, p < .001; $k_{Spelling}$ = 14, r = .40, p < .001). Academic achievement scores on SAT, ACT, or GPA showed a moderate effect size (k = 10, r = .30, p < .001), whereas IQ was related to print exposures with a small effect size (k = 6, r = .18, p = .05). The effect sizes of technical reading and spelling skills and general achievement measures were significantly smaller than the correlation between print exposure and oral language skills. Thus, in line with our second hypothesis, oral language skills were more strongly related to print exposure than technical reading and spelling skills. The correlation between print exposure and reading comprehension outperformed the correlation for basic reading skills (i.e., non-overlapping 95% CIs) but not for word recognition and spelling.

Only one of the moderators that could be tested revealed significant group differences in any of the outcome measures. That is, the effect sizes for students with age-appropriate or higher spelling skills were significantly stronger ($Q(1)$ = 4.86, p < .05; k = 8, r = .45, p < .001) compared to studies that included students with a lower ability (k = 6, r = .29, p < .001). This pattern did not appear to be present for oral language ($Q(1)$ = .19, p > .05).

Reciprocal Causation?

When all age groups are included across meta-analyses, the strength of the correlation between print exposure and oral language showed an increase (see Figure 2.2), whereas the correlations with reading comprehension and technical reading and spelling skills were stable, although they did increase within the set of primary and middle school children. The cross-sectional nature of these studies and variation in spread of scores on skills at different points of mastery, however, stopped us from drawing definite conclusions about print exposure as a consequence of reading ability and as a contributor to further reading growth (i.e., about a causal spiral). The number of longitudinal studies including print exposure checklists was too small to test predictive paths with the meta-analytic approach, but inspection of longitudinal outcomes makes causality more plausible. For children who were followed into elementary school, some researchers did not find predictive relations (e.g., Evans, Shaw, & Bell, 2000; Spear-Swerling, 2006), but others did: For instance, story book exposure in preschool and/or kindergarten significantly explained variance of reading comprehension (6%) and word attack (6%) in first grade but not second grade (Roth, Speece, & Cooper, 2002), reading at the end of third grade (4%; Sénéchal & LeFevre, 2002), and reading comprehension in fourth grade (4%; Sénéchal, 2006). Aram (2005) entered the home literacy environment composite in kindergarten as a first step in predicting second grade skills, explaining 20% of the variance in reading comprehension, 12% in orthographic processing, 16% in spelling, and 12% in text reading fluency, respectively.

Figure 2.2
Effect Size Estimates and 95% CI for Associations between Print Exposure and Oral Language across Years of Education.



Children's own report of print exposure at the end of first grade accounted for 6% of the variance in their third grade reading, after controlling for children's basic reading skills at the beginning of first grade (Sénéchal & LeFevre, 2002). In the same vein, print exposure in third grade contributed to reading comprehension in fifth grade after controlling for third grade reading comprehension (7-11%; Cipeliewski & Stanovich, 1992). Print exposure in fourth to sixth graders explained 8% of oral language and 2% in spelling scores 1.5 years later (Echols, West, Stanovich, & Zehr, 1996). Conversely, two longitudinal studies have shown that print exposure can be predicted by earlier comprehension and technical reading skills. First, reading comprehension and word identification in first grade accounted for 10-12% of the variance in eleventh grade print exposure, as did first grade oral language for 7% and first grade IQ (5% of the variance predicted), after 11^{th}-grade reading comprehension was taken into account (Cunningham & Stanovich, 1997). Third-grade as well as fifth-grade reading comprehension predicted eleventh-grade print exposure as well (22% and 15%, respectively). Second, a variety of basic reading skills, word recognition, and spelling tests in grade 1 and 2 correlated significantly with third grade print exposure, ranging between $r = .40$ and $r = .72$ (Cunningham, Perry, & Stanovich, 2001).

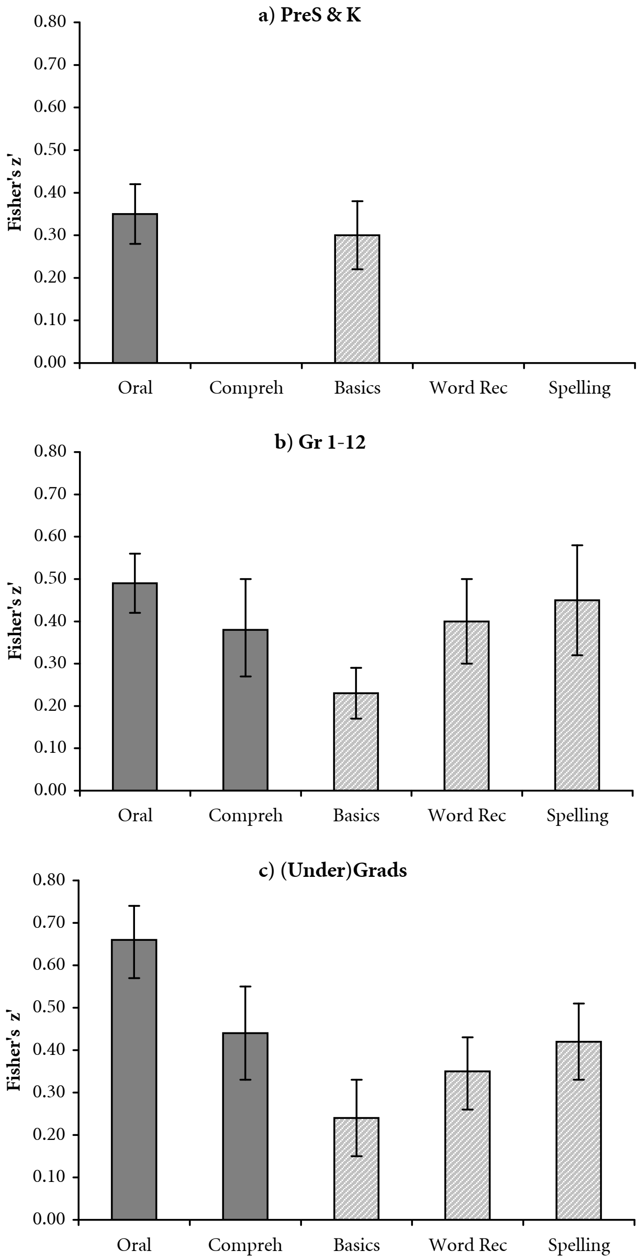
Discussion

We performed a series of meta-analyses on 99 studies ($N = 7{,}669$) that focused on leisure-time reading of preschoolers and kindergartners, children attending grade 1 to 12, and college and university students. The main findings are consistent with a developmental model of reading comprehension and technical reading and spelling, in which print exposure is considered to be a driving force in shaping literacy. In short, it is posited that an early start of shared book reading sets in motion a causal spiral, in which print exposure stimulates language and reading development that, in turn, stimulates the quantity of print exposure (Fletcher & Reese, 2005). For conventional readers, this reciprocal mechanism results in growing inter-individual differences in print exposure that increase with years of education, as more skilled readers will choose to read more and the keener readers will show better comprehension and technical reading and spelling skills (Bast & Reitsma, 1998; Cunningham et al., 1994). Although the meta-analytic results presented herein are largely cross-sectional, precluding a strong stance supporting such a cascading model, the stronger associations between print exposure and several key components of reading skills from infancy to early adulthood are consistent with such a perspective.

Overall, print exposure as inferred from checklists that assess familiarity with book titles and authors or magazines appears to be an important correlate of reading comprehension and technical reading and spelling skill development. During their development, children who choose to read books in their leisure time have larger vocabularies, better reading comprehension, and better technical reading and spelling skills than peers who do not read as frequently. As is displayed in Figure 2.3, the meta-analyses revealed that in the group of 2- to 6-year-old children print exposure is related, at moderate strength, with both oral language and basic reading skills. Second, for children in grades 1 to 12, the moderate effect sizes regarding associations of print exposure with oral language and reading comprehension are comparable to parallel effect sizes found for word recognition and spelling and are significantly stronger than for basic reading skills. Third, the comprehension component (also including academic achievement) and the technical reading and spelling component are moderately to strongly related to print exposure for college and university students, with the effect size for oral language skills the largest of all. In the group of school-aged and university students print exposure is also related to intelligence although effect sizes are small.

Crucially, when we approach our findings from a developmental perspective, the pattern of associations with print exposure was stronger across the age span from early childhood to young adulthood for oral language. Print exposure

Figure 2.3
Effect Sizes for the Comprehension Component (Dark Bars) and Technical Reading and Spelling Component (Lighter Bars) and 95% CI for Studies comprising (a) Preschool and Kindergarten, (b) Grade 1-12, and (c) Undergraduate and Graduate Students.



Note. Oral = Oral Language, *Compreh* = Reading Comprehension, *Basics* = Basic Reading Skills, *Word Rec* = Word Recognition, and *Spelling* = Word Spelling explains 12% of the variance in preschoolers' and kindergartners' oral language skills, 13% in primary school, 19% in middle school, 30% in high school, and 34% at undergraduate and graduate level. The correlation with print exposure also appears to become stronger for technical reading skills and intelligence from primary school to middle school. In addition, print exposure explains significantly more variance in the basic reading skills of school children with low(er) reading abilities (15%) than in their peers with age-appropriate reading abilities (4%). Although these outcomes do not permit conclusions about causality, the pattern of findings as well as a qualitative review of longitudinal studies suggest that spiral causality is a plausible interpretation of our findings.

Book Sharing with Pre-Conventional Readers

In line with the "snowball" metaphor (Raikes et al., 2006), we found that book sharing is associated with not only the development of comprehension but also with technical reading skills that are needed for an easy start at school (see Foster & Miller, 2007). Interestingly, the meta-analysis reveals effects of children's home literacy experiences that are almost identical to those reported in a previous quantitative meta-analysis comprising 33 studies between 1951 and 1993 (Bus et al., 1995). In Bus et al.'s (1995) meta-analysis, the combined effect size was $r = .32$ for oral language and $r = .28$ for reading skills versus $r = .34$ and $r = .29$, respectively in the current data, which covers studies between 1994 and 2008. Even though the earlier meta-analysis included only studies with self-report questionnaires (vs. print-exposure checklists in the current meta-analysis), it is striking that exposure to storybooks explained about 10-12% of children's language and 8% of children's basic reading skills in each investigation. Because effect sizes were comparable for receptive and expressive vocabulary measures, print exposure seems equally effective for language comprehension and language use. Due to insufficient numbers of pertinent studies, we could not test the hypothesis that the association between print exposure and basic reading skills were strongest for kindergartners with more print knowledge, who are more inclined to pay attention to print independently (De Jong & Bus, 2002; Evans et al., 2009).

As oral language and basic reading skills seem to be linked to home environments that familiarize children with books and other reading materials, we see no reason to argue about the recommendation that parents start a reading routine early in children's development. Most longitudinal studies also support the expectation that such a routine prevents pre-conventional readers from experiencing difficulties with understanding print and language in books later on (Aram, 2005; Roth et al., 2002; Sénéchal & LeFevre, 2002; Sénéchal, 2006). The additional finding that parents' knowledge of adult fiction accounts for 7% of children's oral language and basic reading skills is in line with the notion of intergenerational transmission of literacy. That is, if reading is a source of pleasure in their own lives, parents are more inclined to read to their children and engage them in stories (Bus, Leseman, & Keultjes, 2000).

Independent Text Reading

Comprehension. The syntheses of print exposure studies comprising conventional readers revealed moderate to strong effect sizes for oral language and moderate effect sizes for reading comprehension, whereas somewhat smaller effect sizes were found for more distant indicators of the comprehension component such as intelligence and indicators of academic achievement such as GPA and ACT or SAT scores. We argue that a model of reciprocal causation best fits the development of the comprehension component. Developing a reading habit not only depends on environmental factors such as the availability of books at home but also on readers' language and comprehension skills (Stanovich, 1986; Stanovich, Cunningham, & West, 1998). The model predicts that the strength of the correlation between print exposure and language and reading comprehension increases with age, and is strongest for students in college or university who are most likely to be "their own masters" in terms of choosing their leisure time activities.

The comparisons of effect sizes in separate meta-analyses as well as a meta-regression in grade 1 to 12 are consistent with this model of reciprocal causation in particular for oral language. We found a moderate correlation between print exposure and oral language in preschool, kindergarten, primary, and middle school children versus a strong correlation for high school students and undergraduate and graduate students. Impressively, in the development from early childhood to early adulthood, leisure-time reading becomes increasingly more important for language. In early adulthood, 34% of the variance of oral language skills was explained by students' print exposure. We found a similar pattern for intelligence across primary to middle school. Apparently, more intelligent children are more interested in book reading; fiction books cover a huge diversity of topics and thereby provide other perspectives, problems, and/or insights than children might encounter in daily life (Hakemulder, 2000), potentially boosting performance on intelligence tests. More studies are needed, however, that follow children and students longitudinally to learn more about the processes that explain *how* reading might make us "smarter." Apart from the range of cognitive variables as studied in this meta-analysis, future studies should also take into account individual differences in broader cognitive, motivational, socio-emotional, and environmental factors such as general cultural knowledge, interest in reading, skills of empathy and social understanding, and the development of reading routines among other leisure-time activities (e.g., computer use and TV).

We expected that effect sizes for the association between print exposure and reading comprehension would also increase with educational level, because readers' background knowledge expands and their reading strategies get more sophisticated with development (Paris, 2005). However, effect sizes for reading comprehension remained fairly consistent in all age groups. It may be too early to conclude that our findings are in contrast to the model of reciprocal causation, because the comprehension measures seem to have limitations that are likely to influence the effect sizes of print exposure within and across educational levels.

First, reading comprehension tests with relatively brief texts may be easier to complete successfully for older students as compared to younger children, leading to ceiling effects in the oldest age groups that limit the strength of the correlations with print exposure. Second, the expected differences between age groups may not have been captured because the comprehension measures seem to assess different skills in younger and older readers (Cain & Oakhill, 2006; Fletcher, 2006; Keenan, Betjemann, & Olson, 2008). For example, variation in response formats may have masked differences between age groups: Multiple-choice and open-ended questions, which often require integration of text elements, were mainly used in studies on undergraduate and graduate students, whereas relatively easier cloze tasks ("Which alternative word fits best in the sentence?"), which depend more on children's word reading abilities and sentence comprehension, were more often applied in school children. Furthermore, it was impossible to rule out that test scores reflect more general test-taking strategies than reading comprehension (e.g., Ozuru, Rowe, O'Reilly, & McNamara, 2008). Third, most reading comprehension tests may not measure skills that are specific to the comprehension of novels such as following a multi-layered plot and multiple characters throughout hundreds of pages of text as well as understanding complex figures of speech (i.e., metaphors, irony) (Duke, 2010). In contrast, texts in contemporary comprehension tests often comprise brief passages in a variety of genres (e.g., argumentative, expository, narrative) that cover a wide range of topics.

Technical Reading and Spelling. Although instruction is considered to play a main role in learning to read texts with increasing accuracy and fluency (NRP, 2000), the current findings show that print exposure also makes a difference to conventional readers' technical reading and spelling skills. Examining the influence of age in the set of studies on school-aged children, we found that the correlations between print exposure and skills such as basic reading skills, word recognition, and spelling are higher in middle school than in primary school samples, which is in line with reciprocal causality. Readers with higher technical reading and spelling skills are more inclined to read, and more print exposure promotes technical reading and spelling. Even in the studies on college and university students we found that effect sizes for technical reading and spelling skills in relation to print exposure were on the same level. One reason may be that these print exposure studies were conducted in countries with opaque languages such as English, French, and Chinese where children have to familiarize with numerous letter clusters in order to become a skilled reader and where they reach a ceiling in their technical reading and spelling development later than children who learn to read in transparent languages (Furnes & Samuelsson, 2010; Patel et al., 2004; Ziegler & Goswami, 2005). To test this interpretation it will be important to examine technical reading and spelling skills of school children and students who learn to read in more transparent languages (see also Share, 2008). We expect that the technical reading and spelling skills of beginning readers of a language with less extreme ambiguity of spelling-sound correspondences than English will benefit from independent print exposure for a shorter developmental period.

Another reason for the unexpected finding that such associations appear to persist into adulthood may be that outcome measures are constructed in a way that test scores will continue to explain variance in each age group and remain sensitive to differences in students' ability levels even at higher reading proficiency levels. Test adjustments may be made across development to avoid ceiling effects, resulting in unconstrained measures for constrained skills (Paris & Luo, 2010). For instance, the difficulty of words that students must write correctly in a spelling task can be increased for each age group, so that there is enough variance left in the performance of participants to be predicted by print exposure checklists.

In general, a shift occurs in the focus and content of technical reading and spelling measures that are used at different educational levels. For example, alphabet knowledge is only measured in preschoolers, kindergartners, and first graders, which seems methodologically and theoretically sound as no group variance will be left once children received some formal reading instruction and know all letters of the alphabet (Paris, 2005). Phonological and orthographic processing and word recognition appear to be predominantly assessed in children attending primary school, when the most rapid growth in these skills can be expected. By way of contrast, of all the technical reading and spelling skills assessed in college and university students, spelling skills were taken into account most often. It can be argued that at this educational level, variance in reading proficiency may not be effectively captured by a receptive test such as orthographic processing in which correct spellings have to be selected from words that sound similar or by word recognition tasks in which an upper limit may be reached for the speed at which single words can be pronounced. Instead, spelling may be a preferable measure of word-form knowledge because exact knowledge of word forms, especially in English, has to be available in order to write words correctly (Bourassa & Treiman, 2001). As a result of such discrepancies in assessments, direct comparisons of effect sizes for technical reading and spelling skills across age groups may be complicated.

Low-Ability Readers

Leisure time reading is especially important for low-ability readers. We found that the basic reading skills of children in primary and middle school with a lower-ability level were more strongly related to print exposure as compared to higher-ability readers. When low-ability readers have experience with books at home, they practice basic reading skills more, and as a result they become more accurate and fluent in reading text than their low(er)-ability peers who are less exposed to print. The findings suggest that stimulating leisure-time reading should be an effective intervention for low-ability readers as is predicted by the self-teaching hypothesis (Share, 1995). However, for children with reading difficulties it may not be easy to get access to age- and interest-appropriate materials that match their reading ability level and these children may therefore be more dependent on assistance from their parents and/or teachers in selecting stimulating books (Allington & McGill-Franzen, 2008; Kim & White, 2008; Martin et al., 2009).

As for spelling, we found that low-ability readers in studies on college and university students benefited less from print exposure than students whose reading skills fell into the normal range. Older skilled readers may be more capable of deriving word spellings during independent print exposure than less-skilled older readers (Ehri & Saltmarsh, 1995; Reitsma, 1983). We suggest that low-ability readers' uptake of word-specific orthographic details may be limited because they pay attention to words in a text in a way that is qualitatively different from that of more proficient readers. Low-ability readers' use of context information as a compensatory reading strategy may, for instance, interfere with learning word spellings from exposure to print (Ehri & Wilce, 1980; Stanovich, 1986). In all, the current results indicate that encouraging skilled readers to read more may turn them into better spellers, an effect that should not be expected to the same extent for low-ability readers (Nunes & Bryant, 2009; Perfetti & Hart, 2002).

Measurement of Print Exposure

One strength of our meta-analysis is that we were able to compare methods for assessing print exposure by matching studies that administered self-report questionnaire with those utilizing print-exposure checklist studies in the youngest group of pre-conventional readers. A single question about frequency of book reading revealed weaker correlations with oral language and basic reading skills than print-exposure checklists. Such a simple measure is more likely to be positively skewed because it suffers more from (social desirability) biases and therefore shows lower predictive power than the checklist. However, we found no discrepancy between print-exposure checklists and self-report questionnaires when a home literacy composite was used that included a more extensive – and thus more time-consuming – set of questions about the home literacy environment (e.g., the age at which parents started reading, visits to the library and bookstores, number of persons that read to children, parents' ability to mention children's favorite books). The number of books at home – another rather objective indicator of reading volume – reveals effect sizes comparable with print-exposure checklists, further stressing the validity of the checklists as indicators of print exposure.

A relatively small percentage of school-aged children and college and university students completed both a print exposure checklist and a self-report questionnaire about their reading activities or home literacy environment. The moderate to strong correlations between both measurement methods implies that there is overlap in the constructs that are measured by the checklists and questionnaires in these age groups. Interestingly, students who indicated preferring reading as a leisure-time activity to other activities such as listening to music scored higher on print-exposure checklists, whereas students who preferred watching television to reading scored significantly lower on print-exposure checklists. Apparently, print exposure checklists distinguish frequent readers from students who are less likely to choose to read during leisure time. Print-exposure checklists and simply counting books are also less intrusive measures to administer and easier to score than self-report questionnaires. We conclude that checklists and counting books should be preferred as methods to assess print exposure across ages.

Limitations & Future Directions

There are four main limitations of the current meta-analysis. The first is that the findings over-rely on studies conducted in English, whereas different developmental patterns might be found for transparent languages with shallow orthographies.

Second, children from low socioeconomic backgrounds were rarely studied in the youngest age group, probably because researchers expect floor effects on print-exposure checklists in families with limited means and/or few literacy activities. We expect effect sizes in the same range as were detected in our meta-analysis if researchers would succeed to create print exposure checklists that are sensitive to children with varying home literacy experiences. In selecting titles or authors, researchers should take into account that preferences for leisure-time reading materials may vary across socioeconomic status groups and related factors such as ethnicity.

Third, unlike in the set of studies on school children and students in which the same respondent completed the checklist as well as outcome measures, the effect sizes in the youngest group of children were not based on a single respondent. Parents completed the checklists and pre-conventional readers completed the outcome measure(s) which precludes the hypothesis that a third factor such as memory skills or intelligence explains the relation between print exposure and cognitive outcomes (Davidse et al., in press). Interestingly, the effect sizes that are found for primary school children who were administered both a print-exposure checklist and an oral language measure (r = .36) were almost identical to the effect sizes found when parents of somewhat younger children filled in the print-exposure checklists and children completed the language test (r = .34). Therefore, there is not much evidence that the associations merely reflect children's general cognitive capacity.

A fourth limitation is that different measures may have different levels of reliability, which may place constraints on correlations with criterion measures. Larger measurement errors may result in lower correlations (Hunter & Schmidt, 1990). However, in the present set of studies the reliabilities of the measures for print exposure and reading skills were homogeneous and comparably high. For example, the range of reported Cronbach's alpha reliabilities for the print-exposure checklists was between α = .75 and α = .89, which indicates that 75%-89% of the variance is due to the true score and 11%-25% is due to error of measurement. The reliabilities of reading measures were even higher, with alpha reliabilities centering around α =.90. Thus, we do not believe that differential reliabilities were problematic.

Future studies should test the possibility of spiral causality in the reading development of children who are followed longitudinally from infancy through to school age or even adulthood. It would be particularly interesting to identify processes that turn sharing books in infancy into choosing to read as a leisure activity in adolescence and adulthood. For instance, we expect that children's attitudes, beliefs, or motivation towards reading are likely to both influence and depend on current reading skills as well as previous reading experiences, but this has only been examined in a handful of studies so far (e.g., Baker et al., 1997; DeBaryshe, 1995; Guthrie & Wigfield, 1999; Katzir, Lesaux, & Kim, 2009; Kush et al., 2005; Schutte & Malouff, 2007; Shapiro & Whitney, 1997). Knowing why some children choose to read while others do not feel attracted to books might prove useful for the development of successful intervention programs that stimulate skilled as well as less skilled readers of all ages to spend (more of) their leisure time on reading narrative texts.

Conclusions

There is a general belief in society that frequent exposure to print has a long-lasting impact on academic success, as if practicing reading is the miracle drug for the prevention and treatment of reading problems (for reviews, see Dickinson & McCabe, 2001; Phillips, Norris, & Anderson, 2008). This comprehensive meta-analysis of print exposure provides some scientific support for this belief. Our findings are consistent with the theory that reading development starts before formal instruction, with book sharing as one of the facets of a stimulating home literacy environment. Books provide a meaningful context for learning to read, not only as a way of stimulating reading comprehension but also as a means of developing technical reading skills even in early childhood. In pre-conventional readers we found that print exposure was associated moderately with oral language and basic knowledge about reading. Reading books remained important for children in school who were conventional readers. The meta-analyses suggest that reading routines, which are part of the child's leisure-time activities, offer substantial advantages for oral language growth. Interestingly, independent reading of books also enables readers to store specific word form knowledge and become better spellers. Finally, college and university students who read for pleasure may also be more successful academically.

We do not claim that reading more in leisure time is sufficient to turn children into better readers and brighter students in a direct way. Our findings suggest that the relation between print exposure and reading components is reciprocal, as the intensity of print exposure also depends on students' reading proficiency. Print exposure becomes more important for reading components with growing age, in particular for oral language and word recognition. Apparently, children who have developed a reading routine will acquire increasingly more word meanings and word forms from books, which further facilitates their reading development and their willingness to read for pleasure. Such a spiral also implies that readers who lag behind in comprehension or technical reading and spelling skills are especially at risk of developing serious reading problems because they are less inclined to read during leisure time (Stanovich, 1986). With less print exposure, low-ability readers are unlikely to improve their reading and spelling skills to the same extent as their peers who do choose to read. Thus, the reading gap widens and the Matthew effect becomes ever more forceful. Preventing such a downward spiral for poor readers may be among the major challenges of contemporary reading research. We must find ways to motivate these students and their parents to read more as a leisure time activity. In this respect one of our most promising findings is that poor readers' basic reading skills profit most from reading books in their leisure time.

Appendix 2.1

Moderators and Outcomes per Parent-Child Print Exposure Study in Meta-Analysis 1: Preschool and Kindergarten Children.

												Fisher's z (SE)[g]			
First Author[a]	*Year*	*Publ. Status*[b]	*Continent (Country)*	*1st Lang.*[c]	*N*[d]	*% Boys*	*School Type*[e]	*M age (months)*	*Age Range*	*SES*	*Outcome*[f]	*CAR*	*AAR*	*No Bks*	*rFreq*
Aram	2005	Publ.	Asia (Israël)	Oth.	41	46	K	65.59		low	*Oral (EV)*		.18 (.16)	.39 (.16)	
											Basics (PP+OP)	.51 (.09)	.37 (.08)	.46 (.08)	
Burns	2000	Unpubl.	N. Am. (U.S.)	Engl.	59		PreS	36		> middle	*Oral (RV)*	.22 (.13)			
Davidse	2009	Unpubl.	Europe (Neth.)	Oth.	118	52	PreS	54.52	51-57	> middle	*Oral*		.21 (.07)		.14 (.08)
											Basics (OP)		.15 (.10)		.18 (.11)
Evans	2000	Publ.	N. Am. (Canada)	Engl.	66		K	71	65-80	> middle	*Oral (RV)*	.21 (.13)			
											Basics (AK+PP)	.22 (.07)			
Farver	2006	Publ.	N. Am. (U.S.)	Oth.	122	47	PreS	45	39-49	low	*Oral (RV)*	.24 (.09)			.26 (.09)
Foy	2003	Publ.	N. Am. (U.S.)	Engl.	40	43	PreS&K	58.32	48-74.4	> middle	*Oral (EV)*	.06 (.16)			
											Basics (AK+PP)	.09 (.10)			
Frijters	2000	Publ.	N. Am. (Canada)	Engl.	92	54	K	68.5	63-76	> middle	*Oral (RV)*	.41 (.11)		.46 (.11)	.24 (.11)
											Basics (AK+PP)	.30 (.08)		.19 (.08)	.36 (.08)
Gest	2004	Publ.	N. Am. (U.S.)	Engl.	76	59	PreS&K	62.01		low	*Oral*	.56 (.12)			.44 (.12)
											Basics (PP)	.26 (.12)			.29 (.12)

Kalia	2007	Publ.	Asia (India)	Engl.	24		PreS	44.29	40-49	> middle	*Oral*	.52 (.15)			.15 (.15)
											Basics (AK+PP)	.46 (.13)			.15 (.15)
Sénéchal-st1	1996	Publ.	N. Am. (Canada)	Engl.	117	54	PreS&K	52	40-69	> middle	*Oral (RV)*	.45 (.07)	.44 (.09)	.35 (.09)	.24 (.09)
Sénéchal-st2	1996	Publ.	N. Am. (Canada)	Engl.	47	66	PreS&K	49	33-70	> middle	*Oral (RV+EV)*	.45 (.08)	.32 (.11)	.30 (.11)	.19 (.11)
Sénéchal	1998	Publ.	N. Am. (Canada)	Engl.	110	58	K	60	47-79	> middle	*Oral*	.37 (.07)	.25 (.07)		
											Basics (AK+PP+OP)	.30 (.04)	.27 (.04)		
Sénéchal	2000	Publ.	N. Am. (Canada)	Oth.	80	46	K	60	48-71	> middle	*Oral (RV)*	.20 (.08)	.09 (.12)	.17 (.11)	.03 (.11)
Sénéchal	2008	Publ.	N. Am. (Canada)	Engl.	106	46	PreS	56		> middle	*Oral*	.31 (.07)	.29 (.07)		
											Basics (PP)	.35 (.10)	.27 (.10)		
Symons	1996	Publ.	N. Am. (Canada)	Engl.	39	51	K	72		> middle	*Oral (RV)*		.38 (.17)		
											Basics (PP+OP)		.27 (.12)		

Note. [a]*First Author*: For a full list of authors see the reference list (*st1* = study 1, *st2* = study 2, etc. if multiple (independent) samples were derived from the same article); [b]Publication Status: *Publ.* = Published, *Unpubl.* = Unpublished; [c]First Language: *Engl.* = English, *Oth.* = Other; [d]The sample size for the meta-analysis can be smaller based on the data available to calculate Fisher's z'; [e]School Type: *PreS* = PreSchool, *K* = Kindergarten, *PreS&K* = Preschool and Kindergarten children; [f]Outcome: *Oral* = Oral Language Composite, *Oral(EV)* = Oral Language Composite comprises Expressive Vocabulary, *Oral(RV)* = Oral Language Composite comprises Receptive Vocabulary, *Oral(RV+EV)* = Oral Language Composite comprises both Receptive and Expressive Vocabulary Measures, *Basics* = Basic Reading Skills Composite, comprising *AK* = Alphabet Knowledge, *PP* = Phonological Processing, and/or *OP* = Orthographic Processing; [g]Transformed association (Fisher's z (Standard Error)) between *Outcome* and *CAR/CTR* = Child-Author Recognition and Child-Title Recognition Test, *AAR* = Adult-Author Recognition Test, *NoBks* = Number of Books at Home (single item), *rFreq* = Reading Frequency (single item); Only aggregated effect sizes per outcome domain are provided here, see Table 2.2 for weighted combined effect sizes on separate outcome variables.

Appendix 2.2

Moderators and Outcomes for the Matched set of Self-Report Studies in Meta-Analysis 1, in which Each Row Corresponds to the Row of the Print-Exposure Study in Appendix 2.1 that the Self-Report Study is Matched to.

First Author[a]	*Year*	*Publ. Status*[b]	*Continent (Country)*	*1st Lang.*[c]	*N*[c]	*% Boys*	*School Type*[e]	*M age (months)*	*Age Range*	*SES*	*Outcome*[f]	*Fisher's z (SE)*[g] *HLE-Comp*	*Fisher's z (SE)*[g] *rFreq*
Korat	2007	Publ.	Asia (Israël)	Oth.	47	45	K	71.08		low	*Basics*	.24 (.16)	
Deckner	2006	Publ.	N. Am. (U.S.)	Engl.	55	47	PreS	42	42 - 44	> middle	*Oral (RV+EV)*	.47 (.11)	
											Basics (AK)	-.05 (.11)	
Van der Kooy	2009	Unpubl.	Europe (Neth.)	Oth.	101	59	K	64.46	60-72	> middle	*Oral*		.17 (.07)
											Basics (PP)		.17 (.10)
Stephenson	2008	Publ.	N. Am. (Canada)	Engl.	61	49	K	69.84		> middle	*Oral (RV)*	.09 (.13)	.18 (.13)
											Basics (AK+PP)	.18 (.08)	.28 (.09)
Reese	1999	BookCh.	N. Am. (U.S.)	Oth.	121		K			low	*Oral (EV)*	.26 (.13)	
											Basics	.14 (.07)	.08 (.10)
Skibbe	2008	Publ.	N. Am. (U.S.)	Engl.	52	52	PreS	54.02	48-61	> middle	*Basics (AK)*	.26 (.10)	
Roth	2002	Publ.	N. Am. (U.S.)	Engl.	66	58	K	66	62-75	> middle	*Oral*	.41 (.05)	
											Basics (AK+PP)	.22 (.06)	
Roberts	2005	Publ.	N. Am. (U.S.)	Engl.	72	46	PreS	63		low	*Oral (RV+EV)*	.41 (.08)	.20 (.08)
											Basics (AK)	.46 (.13)	.15 (.13)
Kelman	2006	Unpubl. (Diss)	N. Am. (U.S.)	Engl.	91	43	PreS&K	61.36	36-71	> middle	*Oral (RV)*	.21 (.11)	.18 (.11)
											Basics (AK+PP)	.20 (.06)	.18 (.06)

Chaney	1994	Publ.	N. Am. (U.S.)	Engl.	43	51	PreS	44	33-50	> middle	*Oral*	.37 (.08)	
											Basics (AK+PP)	.31 (.15)	
Burgess	2002	Publ.	N. Am. (U.S.)	Engl.	96	52	PreS&K	60.1	48-70	> middle	*Oral (RV+EV)*	.30 (.07)	.14 (.07)
											Basics (AK+PP)	.26 (.03)	
											PP		.17 (.10)
Sénéchal	2006	Publ.	N. Am. (Canada)	Oth.	90	38	K	72		> middle	*Oral (RV)*	.20 (.11)	
											Basics (AK+PP)	.05 (.08)	
Constantine	2005	Unpubl. (Diss)	N. Am. (U.S.)	Engl.	101	56	PreS&K	57.83	48-69	> middle	*Oral*	.29 (.10)	.17 (.10)
											Basics	.21 (.10)	.17 (.10)
Sonnenschein	1996	Unpubl.	N. Am. (U.S.)	Engl.	34		K	70.08		> middle	*Oral (EV)*	.52 (.19)	
											Basics (AK+PP)	.40 (.13)	

Note. [a]*First Author*: For a full list of authors see the reference list; [b]Publication Status: *Publ.* = Published, *Unpubl.* = Unpublished, *Diss* = Dissertation, *BookCh* = Book Chapter; [c]First Language: *Engl.* = English, *Oth.* = Other; [d]The sample size for the meta-analysis can be smaller based on the data available to calculate Fisher's *z*'; [e]School Type: *PreS* = PreSchool, *K* = Kindergarten, *PreS&K* = Preschool and Kindergarten children; [f]Outcome: *Oral* = Oral Language Composite, *Oral(EV)* = Oral Language Composite comprises Expressive Vocabulary, *Oral(RV)* = Oral Language Composite comprises Receptive Vocabulary, *Oral(RV+EV)* = Oral Language Composite comprises both Receptive and Expressive Vocabulary Measures, *Basics* = Basic Reading Skills Composite, comprising *AK* = Alphabet Knowledge, *PP* = Phonological Processing, and/or *OP* = Orthographic Processing; [g]Transformed association (Fisher's *z* (Standard Error)) between *Outcome* and *HLE-Comp* = Home Literacy Environment-Composite, *rFreq* = Reading Frequency (single item); Only aggregated effect sizes per outcome domain are provided here, see Table 2.2 for weighted combined effect sizes on separate outcome variables.

Appendix 2.3

Moderators and Outcomes per Study for Meta-Analysis 2: Grade 1 to 12.

First Author[a]	*Year*	*Publ. Status*[b]	*Continent (Country)*	*Lang. Ch.*[c]	*N*[d]	*% Boys*	*M age (years)*	*Grade*[e]	*Ability Level*[f]	*Check List*[g]	*Fisher's z (SE)*[h] *Oral*	*Compr*	*Basics*	*Word Rec.*	*Spelling*	*IQ*
Allen	1992	Publ.	N. Am. (U.S.)	Engl.	63	60	10.58	5	2	*A+TRT*	.44 (.08)	.57 (.10)				
Barker	1992	Publ.	N. Am. (U.S.)	Engl.	87	49	9.42	3	2	*TRT*			.18 (.09)	.17 (.08)		.09 (.11)
Byrne	1995	Publ.	Australia	Engl.	115	52	8.02	2	2	*TRT*		.20 (.09)		.19 (.07)		
Byrne-st1	1997	BookCh	Australia	Engl.	33			3	1	*TRT*			.39 (.18)			
Byrne-st2	1997	BookCh	Australia	Engl.	28			4	1	*TRT*			.55 (.20)			
Cain	2000	Publ.	Europe (U.K.)	Engl.	25			2	1 vs. 2	*ART*		.17 (.20)				
Cipielewski	1992	Publ.	N. Am. (U.S.)	Engl.	98	53	10.04	4-5	2	*A+TRT*		.50 (.11)	.25 (.08)			
Compton-st1	2002	Publ.	N. Am. (U.S.)	Engl.	32		11.3 vs. 11.6	5-6	1 vs. 2	*TRT*				.66 (.16)		
Compton-st2	2002	Publ.	N. Am. (U.S.)	Engl.	32		8.6 vs. 8.8	3	1 vs. 3	*TRT*				.26 (.17)		
Cunningham	1990	Publ.	N. Am. (U.S.)	Engl.	80	47	9.58	3-4	2	*TRT*			.17 (.21)	.39 (.11)		.17 (.11)
Cunningham	1991	Publ.	N. Am. (U.S.)	Engl.	134	46	11.17	4-6	2	*TRT*	.50 (.09)		.32 (.09)	.59 (.09)	.58 (.09)	.32 (.09)
Cunningham	1993	Publ.	N. Am. (U.S.)	Engl.	26	62	6.92	1	2	*TRT*			.04 (.12)	.59 (.17)	.70 (.15)	
Cunningham	1997	Publ.	N. Am. (U.S.)	Engl.	27	56	16.75	11	2	*A+MRT*	.87 (.23)	.68 (.20)				.32 (.20)
Cunningham	2001	Publ.	N. Am. (U.S.)	Engl.	39	51	9.42	3	1	*TRT*				.59 (.12)		
Ecalle-st1	2008	Publ.	Europe (France)	Oth.	57		6.20	1	2	*A+T +MRT*			.01 (.11)	.21 (.14)	.19 (.14)	-.06 (.14)
Ecalle-st2	2008	Publ.	Europe (France)	Oth.	60		8.34	2-3	2	*A+T +MRT*	.23 (.13)			.37 (.13)	.45 (.13)	
Ecalle-st3	2008	Publ.	Europe (France)	Oth.	60		10.49	4-5	2	*A+T +MRT*	.56 (.13)			.54 (.13)	.39 (.13)	
Echols	1996	Publ.	N. Am. (U.S.)	Engl.	157	54	10.58	4-6	2	*A+TRT*	.49 (.14)				.54 (.06)	
Goff	2005	Publ.	Austr.	Engl.	180		10.17	3-5	2	*TRT*	.39 (.08)	.44 (.08)		.30 (.08)		
Grant-st1	2008	Unpubl. (Conf)	N. Am. (Canada)	Engl.	26	46	8.83	3	1	*TRT*	.17 (.21)	.66 (.21)	.40 (.21)	.56 (.09)		
Grant-st2	2008	Unpubl. (Conf)	N. Am. (Canada)	Engl.	24	42	8.83	3	1	*TRT*	.58 (.22)	.62 (.22)	.29 (.22)	.47 (.10)		
Grant-st3	2008	Unpubl. (Conf)	N. Am. (Canada)	Engl.	18	56	8.63	3	2	*TRT*	.41 (.26)	.14 (.26)	.07 (.26)	-.13 (.12)		

Griffiths	2002	Publ.	Europe (U.K.)	Engl.	59		12.17	large	1	*A+TRT*				.26 (.19)		
Kail	1999	Publ.	N. Am. (U.S.)	Engl.	168	49	10.20	large	2	*A+MRT*		.83 (.08)		.91 (.08)		
Kim	1998	Publ.	Asia (Korea)	Engl.	103	0		high sch.	2	*A+MRT*	.68 (.07)					
Leach-st1	2003	Publ.	N. Am. (U.S.)	Engl.	44	55	10.50	4-5	1 vs. 2	*TRT*		.07 (.15)				
Leach-st2	2003	Publ.	N. Am. (U.S.)	Engl.	60	45	10.50	4-5	1 vs. 2	*TRT*				.17 (.13)		
Lynch	2004	Unpubl. (Diss)	N. Am. (U.S.)	Engl.	56	45		2	2	*TRT*		.23 (.15)		.34 (.15)		
McBride-Chang-st1	1993	Publ.	N. Am. (U.S.)	Engl.	36	69	12.90	5-9	1	*TRT*	.21 (.18)	.76 (.18)	.60 (.18)	.46 (.13)		.27 (.18)
McBride-Chang-st2	1993	Publ.	N. Am. (U.S.)	Engl.	49	49	12.40	5-8	2	*TRT*	.50 (.15)	.33 (.15)	.19 (.15)	.36 (.15)		.39 (.15)
McBride-Chang	1995	Publ.	Asia (China)	Oth.	100	59	11.00	5		*TRT*	.47 (.11)	.30 (.11)				
McBride-Chang	1996	Publ.	N. Am. (U.S.)	Engl.	126	41		large	2	*TRT*	.46 (.09)		.35 (.09)	.56 (.09)		
McDowell	1993	Publ.	N. Am. (U.S.)	Engl.	158	48	8.17	2	2	*TRT*		.10 (.08)	.10 (.08)	.13 (.08)	.14 (.08)	.02 (.08)
McQuillan	2001	Publ.	N. Am. (U.S.)	Engl.	24	42	17.50	11	2	*A+MRT*		.55 (.15)				
McQuillan	2006	Publ.	N. Am. (U.S.)	Engl.	133	52		9-12	1	*ART*	.46 (.09)					
Ricketts	2007	Publ.	Europe (U.K.)	Engl.	30	27	10.07 vs. 9.99	prim. sch.	1 vs. 2	*ART*		.09 (.18)				
Sénéchal	2002	Publ.	N. Am. (Canada)	Engl.	45	51	6.42	1	2	*TRT*	.58 (.15)	.42 (.15)	.26 (.09)	.58 (.11)	.40 (.15)	
Shankweiler	1996	Publ.	N. Am. (U.S.)	Engl.	86		15.17	9	1	*MRT*	.69 (.11)	.44 (.11)	.39 (.11)		.63 (.06)	
Spear-Swerling	2006	Publ.	N. Am. (U.S.)	Engl.	61	48	8.62	3	2	*TRT*		-.09 (.13)				
Stuart	2004	Publ.	Europe (U.K.)	Engl.	53		7.40	2	1	*ART*			.34 (.13)			

Note. [a]*First Author*: For a full list of authors see the reference list (*st1* = study 1, *st2* = study 2, etc. if multiple (independent) samples were derived from the same article); [b]Publication Status: *Publ.* = Published, *Unpubl.* = Unpublished, *Conf* = Conference Contribution, *Diss* = Dissertation; [c]Language Checklist: *Engl.* = English, *Oth.* = Other; [d]The sample size for the meta-analysis can be smaller based on the data available to calculate Fisher's *z*'; [e]Grade: exact grade (range), if not provided *large* = large range, *prim. sch.* = primary school, *high sch.* = high school; [f]Ability Level: *1* = Low(er) Abilities, *2* = Age-Appropriate Abilities, *1 vs .2* = Low(er) Abilities versus Age-Appropriate, *1 vs. 3* = Low(er) versus High Abilities; [g]Checklist: *ART* = Author Recognition Test, *TRT* = Title Recognition Test, *MRT* = Magazine Recognition Test, *A+TRT* = Author and Title Recognition Test, *A+MRT* = Author and Magazine Recognition Test, *A+T+MRT* = Author, Title, and Magazine Recognition Test; [h]Transformed association (Fisher's *z* (Standard Error)) between *Checklist* and: *Oral* = Oral Language Composite, *Compr* = Reading Comprehension, *Basics* = Basic Reading Skills Composite of Alphabet Knowledge, Phonological Processing, and/or Orthographic Processing, *Word Rec.* = Word Recognition Composite of Word Identification and/or Word Attack, *Spelling* = Word Spelling, *IQ* = Intelligence Measure; Only aggregated effect sizes per outcome domain are provided here, see Table 2.3 for weighted combined effect sizes on separate outcome variables.

Appendix 2.4

Moderators and Outcomes per Study for Meta-Analysis 3: Undergraduate and Graduate Students.

First Author[a]	*Year*	*Publ. Status*[b]	*Continent (Country)*	*Lang. Ch.*[c]	*N*[d]	*% Boys*	*M age (years)*	*Student Type*[e]	*Ability Level*[f]	*Check list*[g]	*Fisher's z (SE)*[g] *Oral*	*Compr*	*Basics*	*Word Rec.*	*Spelling*	*IQ*	*Ac Ach.*
Acheson	2008	Publ.	N. Am. (U.S.)	Engl.	99		20.30	Undergr.	2	*ART*							.30 (.08)
										MRT							.03 (.08)
Beech	2002	Publ.	Europe (U.K.)	Engl.	110	18	20.20	Undergr.	2	*ART*	.65 (.07)		.08 (.06)		.46 (.10)		
Burt	2000	Publ.	Australia	Engl.	100	49	19.90	Undergr.	2	*ART*	.46 (.10)	.29 (.10)			.35 (.11)		
Burt	2006	Publ.	Australia	Engl.	112	30	19.80	Undergr.	2	*ART*			.25 (.06)		.44 (.10)		
Chateau	2000	Publ.	N. Am. (Canada)	Engl.	64			Undergr.	2	*ART*			.42 (.13)	.35 (.10)			
Grant-st1	2007	Publ.	N. Am. (Canada)	Engl.	17	24	18.29	Undergr.	2	*ART*	.87 (.27)	.45 (.27)					
										MRT	.18 (.27)	.27 (.27)					
Grant-st2	2007	Publ.	N. Am. (Canada)	Engl.	13	54	21.33	Undergr.	1	*ART*	.56 (.32)	.33 (.32)					
										MRT	.81 (.32)	.79 (.32)					
Hall	1996	Publ.	N. Am. (U.S.)	Engl.	97	18		Undergr.	2	*ART*							.66 (.10)
										MRT							.60 (.10)
Holmes	1993	Publ.	Australia	Engl.	36	19		Undergr.	1 vs. 3	*ART*					.76 (.17)		
Holmes	2001	Publ.	Australia	Engl.	52	17	18.83	Undergr.	1 vs. 3	*ART*					.15 (.14)		
Kennedy	1996	Unpubl. (Diss)	N. Am. (Canada)	Engl.	72	52	20.70	Undergr.	2	*ART*	.76 (.12)	.66 (.12)	.22 (.08)	.32 (.11)			
Krashen	1998	Publ.	N. Am. (U.S.)	Engl.	45	27		Undergr. & Grad.	2	*ART*	1.05 (.15)						
										MRT	.85 (.15)						
Lewellen	1993	Publ.	N. Am. (U.S.)	Engl.	70			Undergr.	1 vs. 2	*ART*	.61 (.26)				.38 (.12)		.59 (.12)
										MRT	.39 (.20)				.13 (.12)		.50 (.12)
Lundquist	2004	Unpubl. (Diss)	N. Am. (U.S.)	Engl.	63	46		Undergr.	1 vs. 3	*ART*	.66 (.13)	.22 (.13)	.22 (.13)	.22 (.13)	.22 (.09)		.27 (.13)
										MRT	.55 (.13)	.22 (.13)	.22 (.13)	.27 (.13)	.22 (.09)		.47 (.13)
Martin-Chang	2008	Publ.	N. Am. (Canada)	Engl.	171	16		Undergr.	2	*ART*	.62 (.08)	.69 (.08)					.16 (.08)
Masterson	2008	Unpubl. (Conf)	Europe (U.K.)	Engl.	80		24.10	Students & Empl.	2	*ART*	.71 (.11)	.28 (.11)			.65 (.11)		
Osana	2007	Publ.	N. Am. (Canada)	Engl.	112	46	23.70	Undergr.	2	*ART*	.79 (.10)	.28 (.10)				-.10 (.10)	.13 (.10)
Rodrigo	1996	Publ.	N. Am.	Oth.	19			Students	2	*ART*	.97 (.25)						

Rodrigo	1996	Publ.	N. Am. (U.S.)	Oth.	19			Students & Empl.	2	*ART*	.97 (.25)						
Sears-st1	2008	Publ.	N. Am. (Canada)	Engl.	75			Undergr. & Grad.	2	*ART*				.21 (.08)			
Sears-st2	2008	Publ.	N. Am. (Canada)	Engl.	76			Undergr. & Grad.	2	*ART*				.48 (.11)			
Siddiqui	1998	Publ.	N. Am. (Canada)	Engl.	133	32	25.80	Undergr. & Grad.	2	*ART* + *MRT*							.38 (.09)
Stanovich-st1	1989	Publ.	N. Am. (U.S.)	Engl.	61	16		Undergr.	2	*ART*					.50 (.13)		
										MRT					.05 (.13)		
Stanovich-st2	1989	Publ.	N. Am.	Engl.	180	37		Undergr.	2	*ART*		.38 (.08)	.32 (.04)	.46 (.06)	.45 (.08)		
										MRT		.30 (.08)	.03 (.04)	.20 (.05)	.11 (.08)		
Stanovich	1992	Publ.	N. Am.	Engl.	300	38		Undergr.	2	*ART*	.73 (.04)	.60 (.06)			.56 (.06)	.31 (.06)	
										MRT	.64 (.04)	.52 (.06)			.39 (.06)	.41 (.06)	
Stanovich	1993	Publ.	N. Am. (U.S.)	Engl.	268	34		Undergr.	2	*ART*		.47 (.06)				.31 (.06)	.34 (.06)
										MRT		.47 (.06)				.30 (.06)	.20 (.06)
Stanovich	1995	Publ.	N. Am.	Engl.	133	28	19.10	Undergr.	2	*ART*	.74 (.09)						.42 (.09)
										MRT	.44 (.09)						.41 (.09)
West	1991	Publ.	N. Am.	Engl.	90	24		Grad.	2	*ART*	.24 (.11)						.03 (.11)
										MRT	.29 (.11)						.10 (.11)
Wolforth-st1	2000	Unpubl. (Diss)	N. Am. (Canada)	Engl.	20	35		Undergr. & Grad.	1	*ART*	.59 (.24)			.16 (.17)	.16 (.24)	.16 (.24)	
										MRT	.16 (.24)			.16 (.17)	.16 (.24)	.16 (.24)	
Wolforth-st2	2000	Unpubl. (Diss)	N. Am. (Canada)	Engl.	21	48		Undergr. & Grad.	1	*ART*	.55 (.24)			.35 (.20)	.16 (.24)	.69 (.24)	
										MRT	.16 (.24)			.29 (.17)	.16 (.24)	.16 (.24)	
Wolforth-st3	2000	Unpubl. (Diss)	N. Am. (Canada)	Engl.	20	40		Undergr. & Grad.	2	*ART*	.16 (.24)			.47 (.27)	.16 (.24)	.16 (.24)	
										MRT	.16 (.24)			.16 (.17)	.16 (.24)	.16 (.24)	

Note. [a]*First Author*: For a full list of authors see the reference list (*st1* = study 1, *st2* = study 2, etc. if multiple (independent) samples were derived from the same article); [b]Publication Status: *Publ.* = Published, *Unpubl.* = Unpublished, *Conf* = Conference Contribution, *Diss* = Dissertation; [c]Language Checklist: *Engl.* = English, *Oth.* = Other; [d]The sample size for the meta-analysis can be smaller based on the data available to calculate Fisher's *z*'; [e]Student Type: *Undergr.* = Undergraduate Students, *Grad.* = Graduate Students, *Empl.* = Employees; [f]Ability Level: *1* = Low(er) Abilities, *2* = Age-Appropriate Abilities, *1 vs. 2* = Low(er) Abilities versus Age-Appropriate, *1 vs. 3* = Low(er) versus High Abilities; [g]Checklist: *ART* = Author Recognition Test, *MRT* = Magazine Recognition Test; [h]Transformed association (Fisher's *z* (Standard Error)) between *Checklist* and: *Oral* = Oral Language Composite, *Compr* = Reading Comprehension, *Basics* = Basic Reading Skills Composite of Phonological and/or Orthographic Processing, *Word Rec.* = Word Recognition Composite of Word Identification and/or Word Attack, *Spelling* = Word Spelling, *IQ* = Intelligence Measure, *Ac Ach.* = *Academic Achievement Indicators* as Grade Point Average and SAT/ACT-scores; Only aggregated effect sizes per outcome domain are provided here, see Table 2.4 for weighted combined effect sizes on separate outcome variables.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

Aarnoutse, C., Van Leeuwe, J., Voeten, M., & Oud, H. (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. *Reading and Writing, 14,* 61-89. Doi: 10.1023/A:1008128417862

* Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40,* 278-289. Doi: 10.3758/BRM.40.1.278

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72,* 167-180. Doi: 10.1177/000312240707200202

* Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology, 84,* 489-503. Doi: 10.1037/0022-0663.84.4.489

Allington, R., & McGill-Franzen, A. (2008). Got books? *Educational Leadership. 65*, 20-23. http://www.ascd.org/publications/educational_leadership/apr08/vol65/num07/Got_Books%C2%A2.aspx

* Aram, D. (2005). Continuity in children's literacy achievements: A longitudinal perspective from kindergarten to school. *First Language, 25,* 259-289. Doi: 10.1177/0142723705050339

Ard, L. M., & Beverly, B. L. (2004). Preschool word learning during joint book reading: Effect of adult questions and comments. *Communication Disorders Quarterly*, *26*, 17-28. Doi: 10.1177/15257401040260010101

Baker, L., Scher, D., & Mackler, K. (1997). Home and family influences on motivations for reading. *Educational Psychologist, 32,* 69-82. Doi: 10.1207/s15326985ep3202_2

Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin*, *129*, 195-215. Doi: 10.1037/0033-2909.129.2.195

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133,* 1-24. Doi: 10.1037/0033-2909.133.1.1

* Barker, T.A., Torgesen, J. K., & Wagner, R. K. (1992). The role of orthographic processing skills on five different reading tasks. *Reading Research Quarterly, 27,* 334-345. Doi: 10.2307/747673

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology, 34,* 1373-1399. Doi: 10.1037/0012-1649.34.6.1373

* Beech, J. R. (2002). Individual differences in mature readers in reading, spelling, and grapheme-phoneme conversion. *Current Psychology, 21,* 121-132. Doi: 10.1007/s12144-002-1008-9

Berninger, V., Abbott, R., Vermeulen, K., Ogier, S., Brooksher, R., Zook, D., et al. (2002). Comparison of faster and slower responders: Implications for the nature and duration of early reading intervention. *Learning Disability Quarterly*, *25*, 59-76. http://www.jstor.org/pss/1511191

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). *Comprehensive meta-analysis version 2.* Biostat, Englewood NJ.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* West Sussex, U.K.: Wiley.

Bourassa, D. C., & Treiman, R. (2001). Spelling development and disability; The importance of linguistic factors. *Language, Speech, and Hearing Services in Schools, 32,* 172-181. Doi:10.1044/0161-1461(2001/016)

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a 'hat' in 'that'? *Journal of Memory and Language, 52*, 131-143. Doi: 10.1016/j.jml.2004.09.003

Bråten, I., Lie, A., Andreassen, R., & Olaussen, B. S. (1999). Leisure time reading and orthographic processes in word recognition among Norwegian third- and fourth-grade students. *Reading & Writing, 11,* 65-88. Doi: 10.1023/A:1007976521114

Breznitz, Z. (1997). Enhancing the reading of dyslexic children by reading acceleration and auditory masking. *Journal of Educational Psychology, 89,* 103-113. http://www.sciencedirect.com/science/journal/00220663

Bulat, J. (2005). *The role of print exposure in the development of early literacy skills among kindergarten students.* Unpublished doctoral dissertation, University of California – Berkeley. (UMI No. 3210520)

* Burgess, S. (2002). The influence of speech perception, oral language ability, the home literacy environment, and pre-reading knowledge on the growth of phonological sensitivity: A one-year longitudinal investigation. *Reading and Writing, 15,* 709-737. Doi: 10.1023/A:1020954606695

Burgess, S. (2005). The preschool home literacy environment provided by teenage mothers. *Early Child Development and Care, 175,* 249-258. Doi: 10.1080/0300443042000266303

* Burns, M. L., & Blewitt, P. (2000). *Parent's book reading: Does quantity of quality predict social class differences in children's vocabulary?* Paper presented at the annual meeting of the Eastern Psychological Association, Baltimore, MD.

* Burt, J. S. (2006). Spelling in adults: The combined influences of language skills and reading experience. *Journal of Pscyholinguistic Research, 35,* 447-470. Doi: 10.1007/s10936-006-9024-9

* Burt, J. S., & Fury, M. B. (2000). Spelling in adults: The role of reading skills and experience. *Reading and Writing, 13,* 1-30. Doi: 10.1023/A:1008071802996

Bus, A. G. (2001). Early book reading in the family: A route to literacy. In: S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research (pp. 179-191).* New York: Guilford Publications.

Bus, A. G. (2003). Social-emotional requisites for learning to read. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 3-15). Mahwah, NJ: Lawrence Erlbaum Associates.

Bus, A. G., & Van IJzendoorn, M. H. (1988). Mother-child interactions, attachment, and emergent literacy: A cross-sectional study. *Child Development, 59,* 1262-1272. Doi: 10.1111/j.1467-8624.1988.tb01495.x

Bus, A. G., Leseman, P. P., & Keultjes, P. (2000). Joint book reading across cultures: A comparison of Surinamese-Dutch, Turkish-Dutch, and Dutch parent-child dyads. *Journal of Literacy Research, 32,* 53-76. Doi: 10.1080/10862960009548064

Bus, A. G., Van IJzendoorn, M. H., & Mol, S. E. (in press). Meta-analysis in reading research. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methods, volume 2.* New York: Guilford Press.

Bus, A. G., Van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1-21. Doi: 10.3102/00346543065001001

* Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology, 87,* 488-503. Doi: 10.1037/0022-0663.87.3.488

* Byrne, B., Fielding-Barnsley, R., Ashley, L., & Larsen, K. (1997). Assessing the child's and the environment's contribution to reading acquisition: What we know and what we don't know. In B. B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia; Implications for early intervention* (pp. 265-285). Mahwah, NJ: Erlbaum.

Cain, K., & Oakhill, J. (2006). Assessment matters: Issues in the measurement of reading comprehension. *British Journal of Educational Psychology, 76,* 697-708. Doi: 10.1348/000709905X69807

* Cain, K., Oakhill, J., & Bryant, P. (2000). Investigating the causes of reading comprehension failure: The comprehension-age match design. *Reading and Writing, 12,* 31-40. Doi: 10.1023/A:1008058319399

Carver, R. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: implications for instruction. *Journal of Reading Behavior, 26*, 413-437. Doi: 10.1080/10862969409547861

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91,* 77-111. Doi: 10.1016/S0010-0277(03)00164-1

Castles, A., & Nation, K. (2008). Learning to be a good orthographic reader. *Journal of Research in Reading, 31,* 1-7. Doi: 10.1111/j.1467-9817.2007.00367.x

Castles, A., Datta, H., Gayan, J., & Olson, R. K. (1999). Varieties of developmental reading disorder: Genetic and environmental influences. *Journal of Experimental Child Psychology, 72,* 73-94. Doi: 10.1006/jecp.1998.2482

Chall, J. S. (1983). *Stages of reading development.* New York: McGraw-Hill.

* Chaney, C. (1994). Language development, metalinguistic awareness, and emergent literacy skills of 3-year-old children in relation to social class. *Applied Psycholinguistics, 15,* 371-394. Doi: 10.1017/S0142716400004501

* Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28,* 143-153. Doi: 35400008197049.0120

Chomsky, C. (1972). Stages in language development and reading exposure. *Harvard Educational Review, 42,* 1-33. http://www.hepg.org/main/her/Index.html

* Cipielewski, J., & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology, 54,* 74-89. Doi: 10.1016/0022-0965(92)90018-2

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Collins, M. F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly, 25,* 84-97. Doi: 10.1016/j.ecresq.2009.07.009

Commission on Reading, National Academy of Education (1985). *Becoming a nation of readers.* Washington, DC: National Institute of Education.

* Compton, D. L. (2002). The relationships among phonological processing, orthographic processing, and lexical development in children with reading disabilities. *Journal of Special Education, 35,* 201-210. Doi: 10.1177/002246690203500402

* Constantine, J. L. (2005). Relationships among early lexical and literacy skills and language-literacy environments at home and school. *Dissertation Abstracts International, 65,* 4456A. (UMI No. 3157208)

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66,* 227-268. Doi: 10.3102/00346543066003227

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal, 110,* 1-18. Doi: 10.1086/598840

* Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing, 14*, 549-568. Doi: 10.1023/A:1011100226798

* Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82,* 733-740. Doi:10.1037/0022-0663.82.4.733

* Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83,* 264-274. Doi:10.1037/0022-0663.83.2.264

* Cunningham, A. E., & Stanovich, K. E. (1993). Children's literacy environments and early word recognition subskills. *Reading and Writing, 5,* 193-204. Doi: 10.1007/BF01027484

* Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934-945. Doi:10.1037/0012-1649.33.6.934

Cunningham, A. E., Stanovich, K. E., & West, R. F. (1994). Literacy environment and the development of children's cognitive skills. In E. M. H. Assink (Ed.), *Literacy acquisition and social context: Approaches, emphases, and questions* (pp. 70-90). London, UK: Harvester Wheatsheaf.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology, 82,* 185-199. Doi: 10.1016/S0022-0965(02)00008-5

Curry, J. E., Parrila, R., Stephenson, K., Kirby, J. R., & Catterson, J. (2004). *The reliability and validity of self-reported home literacy activities and print exposure measures.* Poster presented at the annual meeting of the Society for the Scientific Study of Reading, Amsterdam, The Netherlands.

Daly, A. P. (2000). *Print exposure in eight- to eleven-year-old children: Measurement and prediction.* Unpublished doctoral dissertation, Queen's University of Belfast, United Kingdom.

* Davidse, N. J., De Jong, M. T., Bus, A. G., Huijbregts, S. C. J., & Swaab, H. (in press). Cognitive and environmental predictors of early literacy skills. *Reading and Writing.* Doi: 10.1007/s11145-010-9233-3

DeBaryshe, B. D. (1995). Maternal belief systems: Linchpin in the home reading process. *Journal of Applied Developmental Psychology, 16,* 1-20. Doi:10.1016/0193-3973(95)90013-6

* Deckner, D. F., Adamson, L. B., & Bakeman, R. (2006). Child and maternal contributions to shared reading: Effects on language and literacy development. *Journal of Applied Developmental Psychology, 27,* 31-41. Doi:10.1016/j.appdev.2005.12.001

De Jong, M. T., & Bus, A. G. (2002). Quality of book-reading matters for emergent readers: An experiment with the same book in a regular or electronic format. *Journal of Educational Psychology, 94,* 145-155. Doi: 10.1037//0022-0663.94.1.145

De Jong, P. F., & Leseman, P. P. M. (2001). Lasting effects of home literacy on reading achievement in school. Journal of School Psychology, 39, 389-414. Doi: 10.1016/S0022-4405(01)00080-2

DeTemple, J. & Snow, C.E. (2003). Learning words from books. In A. van Kleeck, S. A. Stahl & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 16-36). Mahwah, NJ: Lawrence Erlbaum.

Dickinson, D. K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research & Practice, 16*, 186-202. Doi: 10.1111/0938-8982.00019

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E.-S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, *95*, 465-481. Doi: 10.1037/0022-0663.95.3.465

Duke, N. (2010, May). *The genre-specific nature of reading comprehension and the case of informational text*. Spoken paper presented at the annual meeting of the AERA, Denver, US.

Duursma, A. E. (2007). Parental bookreading in low-income families: Differences and similarities between fathers and mothers in frequency and style of reading and the impact on children's language and literacy development. *Dissertation Abstracts International, 68,* 2310A. (UMI: No. 3271679).

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method for accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89-98. http://www.jstor.org/stable/2669529

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455-463. Doi: 10.1111/j.0006-341X.2000.00455.x

* Ecalle, J., & Magnan, A. (2008). Relations between print exposure and literacy skills: New evidence from grade 1-5. *British Journal of Developmental Psychology, 26,* 525-544. Doi: 10.1348/026151007X267959

* Echols, L. D., West, R. F., Stanovich, K. E., & Zehr, K. S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology, 88,* 296-304. Doi:10.1037/0022-0663.88.2.296

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word Recognition in Beginning Literacy* (pp. 3-38). Mahwah, NJ: Lawrence Erlbaum.

Ehri, L. C., & Roberts, K. T. (1979). Do beginners learn printed words better in context or in isolation? *Child Development, 50,* 675-685. http://www.jstor.org/pss/1128932

Ehri, L. C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing*, 7, 295-326. Doi: 10.1007/BF03162082

Ehri, L. C., & Wilce, L. S. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology, 71,* 26-40. Doi: 10.1037/0022-0663.71.1.26

Ehri, L. C., & Wilce, L. S. (1980) Do beginners learn to read function words better in sentences or in lists? *Reading Research Quarterly, 15,* 451-76. http://www.jstor.org/pss/747274

Evans, M. A., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science, 16,* 913-920. Doi: 10.1111/j.1467-9280.2005.01636.x

Evans, M. A., & Shaw, D. (2008). Home grown for reading: Parental contributions to young children's emergent literacy and word recognition. *Canadian Psychology, 49,* 89-95. Doi: 10.1037/0708-5591.49.2.89

Evans, M. A., Saint-Aubin, J., & Landry, N. (2009). Letter names and alphabet book reading by senior kindergarteners: An eye movement study. *Child Development, 80,* 1824-1841. Doi: 10.1111/j.1467-8624.2009.01370.x

* Evans, M. A., Shaw, D., & Bell, M. (2000). Home literacy activities and their influence on early literacy skills. *Canadian Journal of Experimental Psychology, 54,* 65-75. Doi: 10.1037/h0087330

* Farver, J. A., Xu, Y., Eppe, S., & Lonigan, C. J. (2006). Home environments and young Latino children's school readiness. *Early Childhood Research Quarterly, 21,* 196-212. Doi: 10.1016/j.ecresq.2006.04.008

Fletcher, J. M.(2006). Measuring reading comprehension. *Scientific Studies of Reading, 10*, 323-330. Doi: 10.1207/s1532799xssr1003_7

Fletcher, K. L., & Reese, E. (2005). Picture book reading with young children: A conceptual framework. *Developmental Review, 25,* 64-103. Doi:10.1016/j.dr.2004.08.009

Forget-Dubois, N., Dionne, G., Lemelin, J-P., Pérusse, D., Tremblay, R. E., & Boivin, M. (2009). Early child language mediates the relation between home environment and school readiness. *Child Development, 80,* 736-749. Doi: 10.1111/j.1467-8624.2009.01294.x

Foster, W.A., & Miller, M. (2007). Development of the literacy achievement gap: A longitudinal study of kindergarten through third grade. *Language, Speech, and Hearing Services in Schools, 38,* 173-181. Doi: 0161-1461/07/3803-0173

* Foy, J. G., & Mann, V. (2003). Home literacy environment and phonological awareness in preschool children: Differential effects for rhyme and phoneme awareness. *Applied Psycholinguistics, 24,* 59-88. Doi: 10.1017/S0142716403000043

* Frijters, J. C., Barron, R. W., & Brunello, M. (2000). Direct and mediated influences of home literacy and literacy interest on prereaders' oral vocabulary and early written language skill. *Journal of Educational Psychology, 92,* 466-477. Doi: 10.1037/0022-0663.92.3.466

Furnes, B., & Samuelsson, S. (2010). Predicting reading and spelling difficulties in transparent and opaque orthographies: A comparison between Scandinavian and US/Australian children. *Dyslexia, 16,* 119-142. Doi: 10.1002/dys.401

* Gest, S. D., Freeman, N. R., Domitrovich, C. E., & Welsh, J. A. (2004). Shared book reading and children's language comprehension skills: The moderating role of parental discipline practices. *Early Childhood Research Quarterly, 19,* 319-336. Doi: 10.1016/j.ecresq.2004.04.007

* Goff, D. A., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing, 18,* 583-616. Doi: 10.1007/s11145-004-7109-0

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *Journal of Memory and Language, 45,* 648-664. Doi: 10.1006/jmla.2001.2790

* Grant, A., Gottardo, A., & Geva, E. (2008). *The mediating role of print exposure in reading comprehension: A comparison between three languages (Spanish, Portuguese, English).* Poster presented at the annual meeting of the Society for the Scientific Study of Reading, Asheville, NC.

* Grant, A., Wilson, A. M., & Gottardo, A. (2007). The role of print exposure in reading skills of postsecondary students with and without reading disabilities. *Exceptionality Education Canada, 17,* 175-194. ERIC#: EJ773919 on http://eric.ed.gov/

* Griffiths, Y. M., & Snowling, M. J. (2002). Predictors of exception word and nonword reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology, 94,* 34-43. Doi:10.1037/0022-0663.94.1.34

Guthrie, J. T., & Wigfield, A. (1999). How motivation fits into a science of reading. *Scientific Studies of Reading, 3,* 199–205. Doi: 10.1207/s1532799xssr0303_1

Hakemulder, J. (2000). *The moral laboratory: Experiments examining the effects of reading literature on social perception and moral self-concept.* Amsterdam, The Netherlands: John Benjamins B.V.

* Hall, V. C., Chiarello, K. S., & Edmenson, B. (1996). Deciding where knowledge comes from depends on where you look. *Journal of Educational Psychology, 88,* 305-313. http://www.sciencedirect.com/science/journal/00220663

Harding, D. W. (1962). Psychological processes in the reading of fiction. *British Journal of Aesthetics, 2,* 133-147. Doi: 10.1093/bjaesthetics/2.2.133

Harlaar, N., Dale, P. S., & Plomin, R. (2007). Reading exposure: A (largely) environmental risk factor with environmentally-mediated effects on reading performance in the primary school years. *Journal of Child Psychology and Psychiatry, 48,* 1192-1199. Doi: 10.1111/j.1469-7610.2007.01798.x

Hayes, D. P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of "motherese"? *Journal of Child Language, 15,* 395-410. Doi:10.1017/S0305000900012411

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11,* 49-76. Doi: 10.1017/S0047404500009039

Hedges, L. V. (2008). What are effect sizes and why do we need them? *Child Development Perspectives, 2,* 167-171. Doi: 10.1111/j.1750-8606.2008.00060.x

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9,* 426-445. Doi: 10.1037/1082-989X.9.4.426

Hirsch, E. D. (2003). Reading comprehension requires knowledge – of words and the word; Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores. *American Educator, 1,* 10-45. Doi: 10.1.1.141.6106

* Holmes, V. M., & Castles, A. E. (2001). Unexpectedly poor spelling in university students. *Scientific Studies of Reading, 5,* 319-350. Doi: 10.1207/S1532799XSSR0504_02

* Holmes, V. M., & Ng, E. (1993). Word-specific knowledge, word-recognition strategies, and spelling ability. *Journal of Memory and Language, 32,* 230-257. Doi: 10.1006/jmla.1993.1013

Hood, M., Conlon, E., & Andrews, G. (2008). Preschool home literacy practices and children's literacy development: A longitudinal analysis. *Journal of Educational Psychology, 100,* 252-271. Doi:10.1037/0022-0663.100.2.252

Hoover, W. A., & Gough, P. B. The simple view of reading. *Reading and Writing, 2,* 127-160. Doi: 10.1007/BF00401799

Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a foreign language, 13,* 403-430. http://nflrc.hawaii.edu/rfl/PastIssues/rfl131hsuehchao.pdf

Hulslander, J., Olson, R. K., Willcutt, E. G., & Wadsworth, S. J. (2010). Longitudinal stability of reading-related skills and their prediction of reading development. *Scientific Studies of Reading, 14,* 111-136. Doi: 10.1080/10888431003604058

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis; Correcting error and bias in research findings*. Newbury Park, CA: Sage.
Jorm, A. F., & Share, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics, 4,* 103-147. Doi: 10.1017/S0142716400004380
Juel, C. (2006). The impact of early school experiences on initial reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research, volume 2* (pp. 410-426). New York: Guilford Press.
Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Developmental Psychology, 44,* 855-866. Doi: 10.1037/0012-1649.44.3.855
Justice, L. M., Skibbe, L., Canning, A., & Lankford, C. (2005). Preschoolers, print and storybooks: An observational study using eye movement analysis. *Journal of Research in Reading, 28,* 229-243. Doi: 10.1111/j.1467-9817.2005.00267.x
* Kail, R., Hall, L. K., & Caskey, B. J. (1999). Processing speed, exposure to print, and naming speed. *Applied Psycholinguistics, 20,* 303-314. Doi:10.1017/S0142716499002076
* Kalia, V. (2007). Assessing the role of book reading practices in Indian bilingual children's English language and literacy development. *Early Childhood Education Journal, 35,* 149-153. Doi: 10.1007/s10643-007-0179-2
Katzir, T., Lesaux, N., & Kim, Y-S (2009). The role of reading self-concept and home literacy practices in fourth grade reading comprehension. *Reading & Writing, 22,* 261-276. Doi: 10.1007/s11145-007-9112-8
Keenan, J. M., Betjemann, R. S., & Olson, R. K.(2008). Reading comprehension tests vary in the skills they assess: differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*, 281-300. Doi: 10.1080/10888430802132279
* Kelman, M. E. (2007). An investigation of preschool children's primary literacy skills. *Dissertation Abstracts International, 68,* 3004B. (UMI No. 3261874)
* Kennedy, A. T. (1996). Individual differences in adult word recognition. *Dissertation Abstracts International, 57,* 735B. (UMI No. NN03584)
* Kim, H., & Krashen, S. (1998). The author recognition and magazine recognition tests, and free voluntary reading as predictors of vocabulary development in English as a foreign language for Korean high school students. *System, 26,* 515-523. Doi: 10.1016/S0346-251X(98)00035-9
Kim, J. S., & White, T. G. (2008). Scaffolding voluntary summer reading for children in grades 3 to 5: An experimental study. *Scientific Studies of Reading, 12,* 1-23. Doi:10.1080/10888430701746849.
Koolstra, C. M., Van der Voort, T. H. A., & Van der Kamp, L. J. T. (1997). Television's impact on children's reading comprehension and decoding skills: A 3-year panel study. *Reading Research Quarterly, 32,* 128-152. Doi: 10.1598/RRQ.32.2.1
* Korat, O., Klein, P., & Segal-Drori, O. (2007). Maternal mediation in book reading, home literacy environment, and children's emergent literacy: A comparison between two social groups. *Reading and Writing, 20,* 361-398. Doi: 10.1007/s11145-006-9034-x
Korat, O., & Schiff, R. (2005). Do children who read more books know "what is good writing" better than children who read less? A comparison between grade levels and SES groups. *Journal of Literacy Research, 37,* 289-324. Doi: 10.1207/s15548430jlr3703_2
Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the Input Hypothesis. *Modern Language Journal, 73,* 440-464. http://www.jstor.org/stable/326879
* Krashen, S., & Kim, H. (1998). The author recognition test without foils as a predictor of vocabulary and cultural literacy test scores. *Perceptual and Motor Skills, 87,* 544-546. Doi: 10.2466/PMS.87.6.544-546
Kush, J. C., Watkins, M. W., & Brookhart, S. M. (2005). The temporal-interactive influence of reading achievement and reading attitude. *Educational Research and Evaluation, 11,* 29-44. Doi: 10.1080/13803610500110141

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon, UK: Multilingual Matters.

* Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95,* 211-224. Doi: 10.1037/0022-0663.95.2.211

Lee, S-Y., & Krashen, S. (1996). Free voluntary reading and writing competence in Taiwanese high school students. *Perceptual and Motor Skills, 83,* 687-690. Doi: 10.2466/PMS.83.6.687-690

Lee, S. Y., Krashen, S., & Tse, L. (1997). The author recognition test and vocabulary knowledge: A replication. *Perceptual and Motor Skills, 85,* 1428-1430. Doi: 10.2466/PMS.85.7.1428-1430

Leppänen, U., Aunola, K., & Nurmi, J. E. (2005). Beginning readers' reading performance and reading habits. *Journal of Research in Reading, 28,* 383-399. Doi: 10.1111/j.1467-9817.2005.00281.x

* Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General, 122,* 316-330. Doi: 10.1037/0096-3445.122.3.316

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Long, D. L., & Prat, C. S. (2002). Memory for Star Trek: The role of prior knowledge in recognition revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1073-1082. Doi: 10.1037/0278-7393.28.6.1073

* Lundquist, E. N. (2004). Phonological complexity, decoding, and text comprehension. *Dissertation Abstracts International, 64,* 6349B. (UMI No. 3116833)

* Lynch, S. C. (2004). A proposed model of reading motivation for young children. *Dissertation Abstracts International, 65,* 411A. (UMI No. 3121589)

Lyytinen, P., Laakso, M.-L., & Poikkeus, A.-M. (1998). Parental contribution to child's early language and interest in books. *European Journal of Psychology of Education, 13,* 297-308. Doi: 10.1007/BF03172946

Mar, R. A. (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation. *Neuropsychologica, 42,* 1414-1434. Doi: 10.1016/j.neuropsychologia.2003.12.016

Mar, R. A., Oatley, K., Hirsh, J., Dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality, 40,* 649-712. Doi:10.1016/j.jrp.2005.08.002

Martin, N. W., Hansell, N. K., Wainwright, M. A., Shekar, S. N., Medland, S. E., Bates, T. C., Burt, J. S., Martin, N. G., & Wright, M. J. (2009). Genetic covariation between the author recognition test and reading and verbal abilities: What can we learn from the analysis of high performance? *Behavioral Genetics, 39,* 417-426. Doi: 10.1007/s10519-009-9275-y

* Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading, 31,* 273-284. Doi: 10.1111/j.1467-9817.2008.00371.x

* Masterson, J., & Hayes, M. (2008). Development and data for UK versions of an author and title recognition test for adults. *Journal of Research in Reading, 30,* 212-219. Doi: 10.1111/j.1467-9817.2006.00320.x

* McBride-Chang, C., & Chang, L. (1995). Memory, print exposure, and metacognition: Components of reading in Chinese children. *International Journal of Psychology, 30,* 607-616. Doi: 10.1080/00207599508246589

* McBride-Chang, C. & Chang, L. (1996). Phonological processing skills and print exposure: Correlates of word reading in children ages 9 to 16. *Reading & Writing Quarterly, 12,* 385-400. Doi: 10.1080/1057356960120403

* McBride-Chang, C., Manis, F. R., Seidenberg, M. S., Custodio, R. G., & Doi, L. M. (1993). Print exposure as a predictor of word reading and reading comprehension in disabled and nondisabled readers. *Journal of Educational Psychology, 85,* 230-238. Doi: 10.1037/0022-0663.85.2.230

McCutchen, D., Harry, D. R., Cunningham, A. E., Cox, S., Sidma, S., & Covill, E. (2002). Reading teachers' knowledge of children's literature and English phonology. *Annals of Dyslexia, 52,* 207-228. Doi: 10.1007/s11881-002-0013-x

McDonald-Connor, C., Son, S. H., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology, 43,* 343-375. Doi:10.1016/j.jsp.2005.06.001

* McDowell, J. A., Schumm, J. S., & Vaughn, S. (1993). Assessing exposure to print: Development of a measure for primary children. *National Reading Conference Yearbook,* 101-107.

McGrath, L. M., Pennington, B. F., Willcutt, E. G., Boada, R., Shriberg, L. D., & Smith, S. D. (2007). Gene x environment interactions in speech sound disorder predict language and preliteracy outcomes. *Development and Psychopathology, 19,* 1047-1072. Doi: 10.1017/S0954579407000533

* McQuillan, J. (2006). The effects of print access and print exposure on English vocabulary acquisition of language minority students. *The Reading Matrix, 6,* 41-51. http://www.readingmatrix.com/articles/mcquillan/article.pdf

* McQuillan, J., & Au, J. (2001). The effect of print access on reading frequency. *Reading Psychology, 22,* 225-248. Doi: 10.1080/027027101753170638

Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64,* 95-114. Doi: 10.1111/j.1540-4560.2008.00550.x

Mistry, R. S., Biesanz, J. C., Chien, N., Howes, C., & Benner, A. D. (2008). Socioeconomic status, parental investments, and the cognitive and behavioral outcomes of low-income children from immigrant and native households. *Early Childhood Research Quarterly, 23,* 193-212. Doi: 10.1016/j.ecresq.2008.01.002

Mol, S. E., Bus, A. G., & De Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research, 79,* 979-1007. Doi: 10.3102/0034654309332561

Mol, S. E., Bus, A. G., De Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education and Development, 19,* 7-26. Doi: 10.1080/10409280701838603

Molfese, V. J., Modglin, A., & Molfese, D. L. (2003). The role of environment in the development of reading skills: A longitudinal study of preschool and school-age measures. *Journal of Learning Disabilities, 36,* 59-67. Doi: 10.1177/00222194030360010701

Moore, D. S., & McCabe, G. P. (2006). *Introduction to the practice of statistics.* New York: W.H. Freeman and Company.

Nagy, W. E. (1988). *Teaching vocabulary to improve reading comprehension*. Newark, DE: International Reading Association.

Nagy, W. E., & Hermann, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Erlbaum.

Nation, K. (2008). Learning to read words. *The Quarterly Journal of Experimental Psychology, 61*, 1121-1133. Doi: 10.1080/17470210802034603

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27,* 342-356. Doi: 10.1111/j.1467-9817.2004.00238.x

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: National Institute of Child Health and Human Development.

Neuman, S. B. (2001). The role of knowledge in early literacy. *Reading Research Quarterly, 36,* 468-475. Doi: 10.1598/ RRQ.36.4.6

Nunes, T., & Bryant, P. (2009). *Children's reading and spelling; Beyond the first steps*. Chichester, U.K.: John Wiley & Sons.

Oatley, K. (1999). Why fiction may be twice as true as fact: Fiction as cognitive and emotional simulation. *Review of General Psychology, 3,* 101-117. Doi: 10.1037/1089-2680.3.2.101

* Osana, H. P., Lacroix, G. L., Tucker, B. J., Idan, E., & Jabbour, G. W. (2007). The impact of print exposure quality and inference construction on syllogistic reasoning. *Journal of Educational Psychology, 99,* 888-902. Doi: 10.1037/0022-0663.99.4.888

Oulette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98,* 554-566. Doi: 10.1037/0022-0663.98.3.554

Ozuru, Y., Rowe, M., O'Reilly, T., & McNamara, D. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods, 40,* 1001-1015. Doi: 10.3758/BRM.40.4.1001

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40,* 184-202. Doi: 10.1598/RRQ.40.2.3

Paris, S. G., & Luo, W. S. (2010). Confounded statistical analyses hinder interpretation of the NELP report. *Educational Researcher, 39,* 316-322. Doi: 10.3102/0013189X10369828

Patel, T. K., Snowling, M. J., & de Jong, P. F. (2004). A cross-linguistic comparison of children learning to read in English and Dutch. *Journal of Educational Psychology, 96,* 785-797. Doi: 10.1037/0022-0663.96.4.785

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., et al. (2000). A cultural effect on brain function. *Nature Neuroscience, 3,* 91-96. Doi: 10.1038/71163

Pavonetti, L. M., Brimmer, K. M., & Cipielewski, J. F. (2003). Accelerated Reader: What are the lasting effects of the reading habits of middle school students exposed to Accelerated Reader in elementary grades? *Journal of Adolescent & Adult Literacy, 46,* 300-311. http://www.jstor.org/pss/40013588

Pecher, D., Zeelenberg, R., & Wagenmakers, E. J. (2005). Enemies and friends in the neighborhood: Orthographic similarity effects in semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 121-128. Doi: 10.1037/0278-7393.31.1.121

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*, 23-33. Doi: 10.1037//0022-0663.94.1.23

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. T. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-214). Amsterdam: John Benjamins Publishing Company.

Petrill, S. A., Deater-Deckard, K., Schatschneider, C., & Davis, C. (2005). Measured environmental influences on early reading: Evidence from an adoption study. *Scientific Studies of Reading, 9,* 237-259. Doi: 10.1207/s1532799xssr0903_4

Phillips, L. M., Norris, S. P., & Anderson, J. (2008). Unlocking the door: Is parents' reading to children the key to early literacy development? *Canadian Psychology, 49,* 82-88. Doi: 10.1037/0708-5591.49.2.82

Radloff, M. E. (2008). Print exposure and performance on the exit level science Texas Assessment of Knowledge and Skills: A correlational study of secondary students. *Dissertation Abstracts International, 69,* 1719A. (UMI No. 3304445)

Raikes, H., Pan, B. A., Luze, G., Tamis-LeMonde, C. S., Brooks-Gunn, J., Constantine, J., Banks Tarullo, L., Raikes, H. A., & Rodriguez, E. T. (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development, 77,* 924-953. Doi: 10.1111/j.1467-8624.2006.00911.x

RAND Reading Study Group (2002). *Reading for understanding: Toward an R&D program in reading comprehension.* Santa Monica, CA: RAND. Doi: 10.1222/0833031058

Rapaport, W. J. (2005). In defense of contextual vocabulary acquisition how to do things with words in context. *Context, 3554,* 396-409. http://www.cse.buffalo.edu/~rapaport/Papers/35540396.pdf

Raviv, T., Kessenich, M., & Morrison, F. J. (2004). A mediational model of the association between socioeconomic status and three-year-old language abilities: The role of parenting factors. *Early Childhood Research Quarterly, 19,* 528-547. Doi: 10.1016/j.ecresq.2004.10.007

* Reese, L., Gallimore, R., & Goldenberg, C. (1999). Job-required literacy, home literacy environments, and school reading: Early literacy experiences of immigrant Latino children. In J. G. Lipson & L. A. McSpadden (Eds.), *Negotiating power and place at the margins: Selected papers on refugees and immigrants* (Vol. 7, pp. 232-269). Arlington, VA: American Anthropological Association.

Reitsma, P. (1983). Word-specific knowledge in beginning reading. *Journal of Research in Reading, 6*, 41-56. Doi: 10.1111/j.1467-9817.1983.tb00237.x

* Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading, 11,* 235-257. Doi: 10.1080/10888430701344306

* Roberts, J., Jurgens, J., & Burchinal, M. (2005). The role of home literacy practices in preschool children's language and emergent literacy skills. *Journal of Speech, Language, and Hearing Research, 48,* 345-359. Doi: 10.1044/1092-4388(2005/024)

* Rodrigo, V., McQuillan, J., & Krashen, S. (1996). Free voluntary reading and vocabulary knowledge in native speakers of Spanish. *Perceptual and Motor Skills, 83,* 648-650. Doi: 10.2466/PMS.83.6.648-650

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (rev. ed.). Newbury Park, CA: Sage.

* Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research, 95,* 259-272. Doi: 10.1080/00220670209596600

Roy-Charland, A., Saint-Aubin, J., & Evans, M. A. (2007). Eye movements in shared book reading with children from kindergarten to grade 4. *Reading and Writing, 20,* 909-931. Doi: 10.1007/s11145-007-9059-9

Sadoski, M.,Willson, V., Holcomb, A., & Boulware-Gooden, R. (2005). Verbal and nonverbal predictors of spelling performance. *Journal of Literacy Research, 36*, 461-478. Doi: 10.1207/s15548430jlr3604_2

Samuelsson, S., Byrne, B., Quain, P., Wadsworth, S., Corley, R., DeFries, J. C., Willcutt, E., & Olson, R. (2005). Environmental and genetic influences on prereading skills in Australia, Scandinavia, and the United States. *Journal of Educational Psychology, 97,* 705-722. Doi: 10.1037/0022-0663.97.4.705

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review, 14,* 245-302. Doi: 10.1006/drev.1994.1010

Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics, 31*, 117-140. Doi: 10.1017/S0142716409990191

Schutte, N. S., & Malouff, J. M. (2007). Dimensions of reading motivation: Development of an adult reading motivation scale. *Reading Psychology, 28,* 469-489. Doi: 10.1080/02702710701568991

* Sears, C. R., Siakaluk, P. D., Chow, V. C., & Buchanan, L. (2008). Is there an effect of print exposure on the word frequency effect and the neighborhood size effect? *Journal of Psycholinguistic Research, 37,* 269-291. Doi: 10.1007/s10936-008-9071-5

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language, 24,* 123-138. Doi: 10.1017/S0305000996003005

* Sénéchal, M. (2000). Examen du lien entre la lecture de livres et le développement du vocabulaire chez l'enfant préscolaire / A naturalistic study of the links between storybook reading and vocabulary development in preschool children. *Enfance, 52,* 169-186. http://www.cairn.info/revue.php?ID_REVUE = ENF

* Sénéchal, M. (2006). Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Scientific Studies of Reading, 10,* 59-87. Doi: 10.1207/s1532799xssr1001_4

Sénéchal, M., & Cornell, E.H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly, 28,* 360-374. http://www.jstor.org/stable/747933

* Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73,* 445-460. Doi: 10.1111/1467-8624.00417

* Sénéchal, M., LeFevre, J. A., Hudson, E., & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of Educational Psychology, 88,* 520-536. Doi: 10.1037/0022-0663.88.3.520

* Sénéchal, M., LeFevre, J., Thomas, E., & Daley, K. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly, 32,* 96-116. Doi: 10.1598/RRQ.33.1.5

* Sénéchal, M., Pagan, S., Lever, R., & Ouellette, G. P. (2008). Relations among the frequency of shared reading and 4-year-old children's vocabulary, morphological and syntax comprehension, and narrative skills. *Early Education and Development, 19,* 27-44. Doi: 10.1080/10409280701838710

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143-174. Doi: 10.1348/000712603321661859

* Shankweiler, D., Lundquist, E., Dreyer, L. G., & Dickinson, C. C. (1996). Reading and spelling difficulties in high school students: Causes and consequences. *Reading and Writing, 8,* 267-294. Doi: 10.1007/BF00420279

Shapiro, J., & Whitney, P. (1997). Factors involved in the leisure reading of upper elementary school students. *Reading Psychology, 18,* 343-370. Doi: 10.1080/0270271970180402

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151-218. Doi: 10.1016/0010-0277(94)00645-2

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72,* 95-129. Doi: 10.1006/jecp.1998.2481

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584-615. Doi: 10.1037/0033-2909.134.4.584

Share, D. L., & Shalev, C. (2004). Self-teaching in normal and disabled readers. *Reading and Writing, 17,* 769-800. Doi: 10.1007/s11145-004-2658-9

Shatil, E., & Share, D. L. (2003). Cognitive antecedents of early reading ability: A test of the modularity hypothesis. *Journal of Experimental Child Psychology, 86,* 1-31. Doi: 10.1016/S0022-0965(03)00106-1

Shaywitz, S. E., & Shaywitz, B. A. (2008). Paying attention to reading: The neurobiology of reading and dyslexia. *Development and Psychopathology*, *20*, 1329-1349. Doi: 10.1017/S0954579408000631

* Siddiqui, S., West, R. F., & Stanovich, K. E. (1998). The influence of print exposure on syllogistic reasoning and knowledge of mental-state verbs. *Scientific Studies of Reading, 2,* 81-96. Doi: 10.1207/s1532799xssr0201_4

* Skibbe, L. E., Justice, L. M., Zucker, T. A., & McGinty, A. S. (2008). Relations among maternal literacy beliefs, home literacy practices, and the emergent literacy skills of preschoolers with specific language Impairment. *Early Education & Development, 19,* 68-88. Doi: 10.1080/10409280701839015

* Sonnenschein, S., Baker, L., Serpell, R., Scher, D., Fernandez-Fein, S., & Munstermann, K. A. (1996). *Strands of emergent literacy and their antecedents at home: Urban preschoolers' early literacy development* (Reading Research Report No. 48). Georgia & Maryland: National Reading Research Center.

* Spear-Swerling, L. (2006). Children's reading comprehension and oral reading fluency in easy text. *Reading and Writing, 19,* 199-220. Doi: 10.1007/s11145-005-4114-x

Stainthorp, R. (1997). A Children's Author Recognition Test: A useful tool in reading research. *Journal of Research in Reading, 20,* 148-158. Doi: 10.1111/1467-9817.00027

Stainthorp, R., & Hughes, D. (1998). Family literacy activities in the homes of successful young readers. *Journal of Research in Reading, 23,* 41-54. Doi: 10.1111/1467-9817.00101

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-407. Doi: 10.1598/RRQ.21.4.1

Stanovich, K. E. (2000). Measuring print exposure; Attempts to empirically track "rich get richer" effects. In K. E. Stanovich, *Progress in understanding reading; Scientific foundations and new frontiers* (pp. 245-258). New York: The Guilford Press.

* Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition, 20*, 51-68. http://www.psychonomic.org/backissues/5413/mc/vol20-1/

* Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology, 85,* 211-229. Doi:10.1037/0022-0663.85.2.211

Stanovich, K. E., Cunningham, A. E., & West, R. F. (1998). Literacy experiences and the shaping of cognition. In S. G. Paris & H. M. Wellman (Eds.), *Global prospects for education: Development, culture, and schooling* (pp. 253-288). Washington, DC: APA.

* Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24,* 402-433. Doi: 10.2307/747605

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89,* 114-127. Doi: 10.1037/0022-0663.89.1.114

* Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology, 31,* 811-826. Doi:10.1037/0012-1649.31.5.811

* Stephenson, K. A., Parrila, R. K., Georgiou, G. K., & Kirby, J. R. (2008). Effects of home literacy, parents' beliefs, and children's task-focused behavior on emergent literacy and word reading skills. *Scientific Studies of Reading, 12,* 24-50. Doi: 10.1080/10888430701746864

Stone, D. B., Fisher, S. K., & Eliot, J. (1999). Adults' prior exposure to print as a predictor of the legibility of text on paper and laptop computer. *Reading and Writing, 11,* 1-28. Doi: 10.1023/A:1007993603998

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, *38*, 934–947. Doi: 10.1037//0012-1649.38.6.934

* Stuart, M. (2004). Getting ready for reading: A follow-up study of inner city second language learners at the end of Key Stage 1. *British Journal of Educational Psychology, 74,* 15-36. Doi: 10.1348/000709904322848806

Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. *Reading Research Quarterly, 20,* 458-481. Doi: 10.1598/RRQ.20.4.4

Swanborn, M. S. L., & De Glopper, K. (2002). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69,* 261-285. Doi: 10.3102/00346543069003261

* Symons, S., Bonnell, C., & Szusziewicz, T. (1996). *Parental Print Exposure and Young Children's Language and Literacy Skills*. Retrieved April 22, 2008, from http://www.nald.ca/library/research/pat/march96/page49.htm

Teale, W. H. (1981). Parents reading to their children: What we know and need to know. *Language Arts, 58,* 902-911. ERIC#: EJ254985 on http://eric.ed.gov/

Teale, W. H., & Sulzby, E. (1986). *Emergent Literacy: Writing and Reading.* Norwood, NJ: Ablex.

* Van der Kooy, V. A. C., Kegel, C. A. T., & Bus, A. G. (in press). Evidence-based computer interventions targeting phonological awareness to prevent reading problems in at-risk young students. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research, volume 3* (pp. 214-227). New York: Guilford Press.

Van Lierop-Debrauer, H. (1990). *Ik heb het wel in jouw stem gehoord* [I did hear it in your voice]. Delft: Eburon.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11,* 3-32. Doi: 10.1080/10888430709336632

* West, R. F., & Stanovich, K. E. (1991). The incidental acquisition of information from reading. *Psychological Science, 1991,* 325-330. Doi: 10.1111/j.1467-9280.1991.tb00160.x

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly, 28,* 34-50. Doi: 10.2307/747815

* Wolforth, J. B. (2000). Reading and listening comprehension in university students with and without reading disability. *Dissertation Abstracts International, 60,* 2858A. (UMI No. NQ41082)

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., & Fischel, J. E. et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24,* 552-559. Doi: 10.1037/0012-1649.24.4.552

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848-872. Doi: 10.2307/1132208

Wood, C. (2002). Parent-child pre-school activities can affect the development of literacy skills. *Journal of Research in Reading, 25,* 241-258. Doi: 10.1111/1467-9817.00173

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3-29. Doi: 10.1037/0033-2909.131.1.3

3

Added Value of Dialogic Parent-Child Book Readings: A Meta-Analysis

Abstract

Book reading has been demonstrated to promote vocabulary. The current study was conducted to examine the added value of an interactive shared book reading format that emphasizes active as opposed to non-interactive participation by the child. Studies that included a Dialogic Reading intervention group and a reading-as-usual control group, and that reported vocabulary as an outcome measure were located. After extracting relevant data from 16 eligible studies, a meta-analysis was conducted to attain an overall mean effect size reflecting the success of Dialogic Reading in increasing children's vocabulary compared to typical shared reading. When focusing on measures of expressive vocabulary in particular ($k = 9$; $n = 322$), Cohen's d was .59 ($SE = .08$; *95%* CI = .44, .75, $p < .001$), which is a moderate effect size. However, the effect size reduced substantially when children were older (five- to six-years-old) or when they were at risk for language and literacy impairments. Dialogic Reading can change the home literacy activities of families with two- to four-year-old children but not those of families with children at greatest risk for school failure.

Based on:
Mol, S. E., Bus, A. G., De Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education and Development, 19,* 7-26. Doi: 10.1080/10409280701838603

Introduction

Exposure to books is a major source for the development of one of the main areas in learning to read: vocabulary (Bus, van IJzendoorn, & Pellegrini, 1995). The importance of vocabulary appears from findings indicating that the vocabulary of children entering first grade predicts not only their word reading ability at the end of first grade but also reading comprehension later in their school career (Juel, 2006). Books stimulate vocabulary because they contain a wider vocabulary than occurs in ordinary conversations (Sulzby, 1985). Even simple stories for two- and three-year-olds such as the Maisy stories (by Lucy Cousins) include complex words and phrases such as 'starving' or 'right on time' with a much higher incidence in books than in daily communication. In particular as a result of repeated encounters with the same story, young children appear to expand their vocabulary (e.g., Verhallen, Bus, & de Jong, 2006).

Assuming a social-constructionist nature of book reading, books cannot be a source for acquiring new vocabulary unless children get intensive help and support from adults. Consequently, children may almost never encounter solely an oral rendering of the text. Instead, in most cases the words of the author are accompanied by the social interaction between the adult reader and child. The book reading paradigm assumes that whether or not children become interested in books and learn new vocabulary through shared book reading depends on this social context. Following this reasoning, parent-child storybook reading is hypothesized to be effective mainly as a result of tutorial behavior from mothers or fathers (Ninio & Bruner, 1978). It is assumed that children benefit most from book reading designed to increase child responding at the expense of simply listening to parents read (e.g., DeTemple & Snow, 2003; Raikes et al., 2006). That is, book reading may be most effective when parents seek to involve the child actively in verbal exchanges when reading aloud (Huebner & Meltzoff, 2005). In line with this hypothesis, a seminal study by Whitehurst and colleagues (Whitehurst, Falco, Lonigan, et al., 1988) followed by a series of replications in groups of different age and socioeconomic status (SES) experimentally tested the relations between adult tutoring during book reading and children's linguistic development. Assuming that meta-analyses can be applied most fruitfully within research programs in which studies with similar designs (and measures) have accumulated over the years (Bus & van IJzendoorn, 2004), we decided to take a meta-analytic approach to test the tenability of Whitehurst's conclusion that variations in reading to young children can have appreciable effects on language development.

When parents approach storybook reading with the intent of teaching language to their children, they may use techniques such as asking questions, giving feedback, and adjusting questions to the developmental level of the child. Because observational studies have suggested that most parents do not apply interactive reading techniques spontaneously (e.g., Britto, Brooks-Gunn, &

Griffin, 2006; Laakso, Poikkeus, & Lyytinen, 1999; Silvén, Ahtola, & Niemi, 2003), it seems reasonable that parents should receive training in this. Whitehurst et al. (1988) were the first to report promising results on vocabulary as a result of an intervention based on a stimulation package called "Dialogic Reading", which was designed according to the following three principles: (a) the use of evocative techniques by the parent that encourage the child to talk about pictured materials; (b) informative feedback by incorporating expansions, corrective modeling, and other forms that highlight differences between what the child has said and what he might have said; and (c) an adaptive parent sensitive to the child's developing abilities. In sum, during typical shared reading the adult reads and the child listens, but in Dialogic Reading the child learns to become the storyteller; Dialogic Reading involves reading with rather than to children (Fielding-Barnsley & Purdie, 2003). A positive demonstration of this intervention could set the stage for more fine-grained analyses of particular processes mediating effects, but so far most follow-up intervention studies include this broad type of Dialogic Reading. Because replications of Whitehurst's influential study also have included children that have differed in age and risk status from the participants in Whitehurst's study, the set of studies enabled us to test which variables may moderate outcomes of the intervention.

Using an exhaustive set of studies investigating the effects of Dialogic Reading, we tested whether variations in parental reading affect children's language development. The experimental groups participated in a treatment program that instructed parents to alter the frequency and timing of various aspects of their child-directed speech during story time. Control families were assumed to read to their children but were not told to change any behaviors. Based on this set of studies, we aimed to test whether there is evidence for the hypothesis that Dialogic Reading causes effects beyond those of the typical format of book reading. At the time of the first meta-analysis of parent-child storybook reading (Bus et al., 1995), there were insufficient studies available to include dialogue between parents and children as a moderator variable and test the additional effects of the variations in book reading described above on learning outcomes. The frequency of parent-child storybook reading explained no less than 8% of the variance in language and literacy measures (Bus et al., 1995). By treating research group as a moderator variable, we tested whether the strong effect sizes found in the pioneering studies by Whitehurst and his colleagues could be replicated by subsequent research (Bus & Van IJzendoorn, 2004).

Focusing on vocabulary gains, the main indicator of results in all studies, we expected to find stronger effects for expressive than receptive vocabulary. That is, Dialogic Reading is particularly known for its role in stimulating active verbal involvement by the child and is therefore especially thought to support expressive language. We also expected Dialogic Reading to be more effective in younger age groups as their active participation and learning are highly dependent on the parental ability to bridge the discrepancy between the child's world and the world of the book. Older and more experienced children may rely less on external support in understanding and enjoying the story and may be more inclined to initiate dialogue when they do not grasp the story content.

Lastly, we expected that studies including more low-educated families would reveal less pronounced effects than studies including higher educated families. Whitehurst et al.'s (1988) study focused on intact middle-class families living on suburban Long Island, New York. In line with Ninio (1980), who examined social class differences when mothers read to two-year-old children, it is to be expected that low-educated mothers may be less likely to engage in a number of potentially instructive behaviors during story time. In the same vein, Heath (1982) found that high-SES mothers pose more "why" questions than low-SES mothers do. Bus and van IJzendoorn (1995) observed that a high-SES mother is inclined to guide her three-year-old child's understanding of the story plot by asking questions and offering help to answer questions, whereas the low-SES mother, who reads infrequently, just explains parts of the story without any attempt to involve the child in thinking about the event. There is also evidence that when adults do not find reading books a source of pleasure for themselves, then activities such as book reading may not be embedded in family practice, and parents may not know how to engage children in reading sessions (Bus, Leseman, & Keultjes, 2000). In the context of these findings we wonder whether the technique of Dialogic Reading works for both children developing typically and those children at risk for language and literacy impairments (Hargrave & Sénéchal, 2000).

To sum up, this meta-analysis tested the following hypotheses:

1) Does Dialogic Reading intensify the effects of parent-child picture storybook sharing, and how strong is the additional effect of Dialogic Reading? We expected that Dialogic Reading would add to the effects of typical book reading.
2) Does Dialogic Reading affect expressive language skills more strongly than receptive language skills? Expressive language skills may be particularly affected because those skills are emphasized within the Dialogic Reading format.
3) Is the strength of the association between Dialogic Reading and outcome measures related to the age at which the intervention started? We hypothesized that older children with more linguistic skills would be less dependent on the qualities of book reading sessions than younger children with inferior linguistic and comprehension skills.
4) Is Dialogic Reading as strongly related to outcome measures in samples of young children at risk for language and literacy impairments as in samples of children not at risk? It was hypothesized, for instance, that at-risk parents may be less responsive to training in book-sharing skills because reading is not a source of pleasure for many parents in this group.

5) Were stronger effect sizes revealed by the group who started this line of research than other researchers replicating the study, as is apparent in other research domains (Bus & van IJzendoorn, 2004)? If this is the case, studies carried out by Whitehurst and colleagues should be more effective than studies carried by other research groups.

Method

Search Strategy

An extensive literature search was conducted across Psychological Abstracts Online (PsycINFO), the Education Resources Information Center (ERIC), Dissertations Abstracts Online and SilverPlatter's Information Retrieval System for the Web (WebSPIRS) to identify all eligible studies from 1988 to March 2007. The following keywords were used during this computer search: *(shared / interactive / dialogic) book reading, (early) intervention, reading intervention, home literacy environment, parent, parent-child interaction, achievement, language development, vocabulary, preschool, kindergarten,* and *young children*. Additional studies were identified by manual search in reference lists of previous meta-analyses and primary studies. To retrieve unpublished documents or clarify uncertainties, several authors were contacted.

Selection Criteria

To be included in the present meta-analysis, studies had to describe original data and meet the following criteria: (a) involve Dialogic Reading programs in which parents were trained to read interactively with their child; (b) include pre-conventional reading participants with no mental, physical, or sensory handicaps and are pre-conventional readers; (c) contain outcome variables that were objective measures of expressive and/or receptive vocabulary; (d) involve a (quasi)experimental design, that included a control group in which parents were asked to read as usual; (e) be reported in English (although no restriction was put on the country and therefore the language in which the study was conducted); and (f) be published or unpublished.

We excluded studies if the intervention involved teacher- and/or stranger-child rather than parent-child shared book reading (e.g., Valdez-Menchaca & Whitehurst, 1992; Wasik & Bond, 2001), or if the intervention was a combined school- and home-based program in which no separate data for parent-child interaction were reported (Hargrave & Sénéchal, 2000; Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999; Whitehurst, Arnold, et al., 1994; Whitehurst, Epstein, et al., 1994; Whitehurst et al., 1999). Studies were also excluded when the focal children suffered from physical handicaps such as hearing impairment (Fung, Chow, & McBride-Chang, 2005). Furthermore, one study was eliminated from the initial sample because children in the control group received a language-related intervention (Dale, Crain-Thoreson, Notari-Syverson, & Cole, 1996). Finally, some studies were excluded because no data were presented (e.g., Gormley & Ruhl, 2005; Whitehurst & Lonigan, 1998; Zevenbergen & Whitehurst, 2003) or a control group was missing (Huebner, 2000a, Huebner & Meltzoff, 2005). Unfortunately, one relevant study (Park, 2006) focusing on 20 Korean, two- to four-year-old children learning English in the United States, could not be located. The abstract suggests that a positive but not significant effect for the Dialogic Reading group was present in this study.

Coding of the Studies

We coded all relevant studies using a standardized data extraction tool into which the following study characteristics could be entered: (a) *bibliographic reference*: publication type, year of publication; (b) *sample descriptors*: number of participants per group, mean age and school type at the start of the study (preschool or kindergarten), country of origin (Asia, Australia, Europe, the United States), language(s) used in the intervention, risk status (at risk, not at risk); (c) *research design descriptors:* design (experimental, quasi-experimental), type of Dialogic Reading training (video instruction, group session, individual training), duration of the intervention (in weeks), check of treatment integrity (use of Dialogic Reading techniques and the frequency of book reading in the experimental and the control group; scale = 0-8); and (d) *outcome measures:* test(s) used to measure vocabulary (receptive and/or expressive), posttest data (mean and standard deviation, *t*-test, *F*-test, *p*-value, sample size per test).

As indicators of expressive vocabulary, we included the Expressive One-Word Picture Vocabulary Test (EOWPVT; Gardner, 1981), the expressive vocabulary subtest of the Illinois Test of Psycholinguistic Abilities (ITPA-VE; Kirk, McCarthy, & Kirk, 1968), and mean and/or total length of utterances by the child during videotaped reading sessions. As measures for receptive vocabulary we used the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) and the Bracken Basic Concept Scale (BBCS; Bracken, 1984), which assesses conceptual development and tests concepts such as colors, materials, positions, sequences, and shapes.

If studies included more than one intervention or control group and all groups met the inclusion criteria, sample sizes were adapted. For example, Arnold, Lonigan, Whitehurst, and Epstein (1994) trained parents in the intervention group either by showing them a video or via group sessions. In this study, the sample size of the control group was split into two groups in order to include results from both Dialogic Reading-interventions. Chow, McBride-Chang, Cheung, and Chow (2007) included a typical reading control group and a no intervention control group in addition to an experimental group comprising parents trained to read dialogically. In the typical reading group, parents kept a log of their reading frequency and were asked to read as usual with their child when they received the same picture storybooks as the parents in the Dialogic Reading group, whereas the children in the control group were only pre- and posttested. Because the intervention group could not be included twice, the sample size was divided by two. The means and standard deviations remained unchanged.

To assess inter-coder reliability, two coders independently coded all selected studies. Both coders agreed completely on including the same list of studies. The average percentage of agreement across study characteristics and moderators was 96% (κ = .94, *range* = .60 – 1.00). Discrepancies between coders were resolved by discussion until consensus was reached.

The coding process resulted in a final set of sixteen studies, of which eight studies reported measures of both children's receptive and expressive vocabulary. Of the remaining studies, seven tested only receptive vocabulary, whereas one focussed solely on expressive vocabulary. A total of 626 parent-child dyads ($N_{DR\text{-}intervention}$ = 313; $N_{Control}$ = 313) were studied. Participants' mean age ranged from 27.8 to 70.2 months. For further details about the characteristics of the studies included, see Appendix 3.1.

Meta-Analytic Procedures

The standardized difference between the mean of a Dialogic Reading-intervention group and a reading-as-usual control group at posttest was computed in order to quantify the additional value of Dialogic Reading on vocabulary. Because authors presented various statistics, we used Wilson's Effect Size Calculator (Wilson, 2001) and Comprehensive Meta-Analysis software (Version 2.2; Borenstein, Hedges, Higgins, & Rothstein, 2005) to calculate Cohen's *d* effect sizes for each vocabulary outcome. A positive effect size indicates a favorable outcome for the Dialogic Reading intervention group. A *d* of .20 is interpreted as a small, .50 as a moderate, and .80 as a large effect size (Cohen, 1992). Because studies with an increased sample size provide more reliable estimates of the population mean due to a smaller standard error, effect sizes were determined by weighting each outcome by the inverse of its variance (Cooper & Hedges, 1994; Lipsey & Wilson, 2001). However, some studies contained more than one outcome measure or effect size. To prevent unequal weighting of effect sizes due to the number of measurements in the study (e.g., some reported one receptive and two expressive vocabulary measures), effect sizes were aggregated within a domain before being averaged across outcomes within a study. Then, an overall combined Cohen's *d* was calculated per study before further analyses were conducted.

Bias due to the fact that studies with non-significant findings are less likely to be published was examined graphically by funnel plot analysis. For each study, effect sizes were plotted against precision as determined by the sample size or the inversed standard error, to detect a potential bias due to under-representation of studies with small sample sizes. We used the "trim and fill" method to calculate the effect of potential file drawer problems (Duval & Tweedie, 2000a, 2000b). We also computed the fail-safe number (*N*fs), i.e., the number of studies with null results that have not been published but have to exist to overturn the association between Dialogic Reading and vocabulary gains to a level of non significance (Lipsey & Wilson, 2001). Furthermore, homogeneity across studies was assessed by means of the *Q*-statistic to determine whether variability among individual effect sizes was larger than should be expected based on subject level sampling error. Significant *Q*s imply heterogeneity, which indicates that the separate effect sizes do not all estimate the same population mean effect size (Lipsey & Wilson, 2001). Additionally, I-squared (I^2) is presented, which measures the degree of inconsistency between studies. I^2-values larger than 75% implying heterogeneity (Petticrew & Roberts, 2006). When heterogeneity was indicated by both measures, a random effects model was preferred, assuming that variability between studies was random instead of systematic (Rosenthal, 1995).

Effects of moderator variables such as population at school, risk status, publication year, and treatment integrity were tested by contrasting sub-samples or by applying a meta-regression model. It is important to note that each subset had to consist of at least four studies before contrasts were considered (Bakermans-Kranenburg, van IJzendoorn, & Juffer, 2003). Fixed effects models were applied in the case of homogeneous sets of outcomes, either by a significant $Q_{between}$ (heterogeneity between subsets) and a non-significant Q_{within} (homogeneity within the subsets), or an I^2 less than 75%. More conservative random effects model tests were applied when heterogeneous outcomes were present.

Results

Preliminary Analysis

No outlying values appeared to be present. That is, standardized *z* values were smaller than 3.26 and/or larger than -3.26 for all effect sizes ($p < .001$). Evidence for publication bias was absent, as the funnel plot of precision showed symmetry around the point estimate. Further, we tested whether more controlled experiments revealed stronger results than studies that hardly checked the content or frequency of book reading sessions. A meta-regression showed no significant differences ($d = .05$, $p > .05$), implying that the quality of the intervention did not affect the overall effect size. Furthermore, a moderator analysis revealed that studies conducted within the research group of Whitehurst ($k = 5$ studies, $n = 136$ children) did not significantly differ from experiments by other research groups ($Q = 3.64$; $p > .05$). The minor variation in duration of the interventions did not influence any effects either. We were not able to test for publication bias by comparing published and unpublished studies because only two unpublished studies were located.

The fixed effects model was applied in order to calculate the mean effect size for overall, expressive, and receptive vocabulary because either *Q* or I^2 indicated that the samples were homogeneous ($Q_{overall} = 34.10$, $p < .01$, $I^2 = 56.01$; $Q_{receptive} = 21.34$,

$p > .05$, $I^2 = 34.40$; $Q_{expressive} = 11.27$, $p > .05$, $I^2 = 29.00$). Standardized differences in means, 95% Confidence Intervals (CIs) per study, and outcome measure are presented in Appendix 3.1.

The Additional Effect of Dialogic Reading Intervention on Vocabulary Measures

For all included studies ($k = 16$, $N = 626$), Cohen's d equaled .42 ($SE = .06$; *95%* CI = .30, .53; $p < .001$). The fail-safe number indicated that 123 additional studies with null or non-significant results needed to be added to negatively influence this significant but small effect size. An overview of the distribution of the combined effect sizes per study is graphically displayed in Table 3.1.

Table 3.1

Stem-and-Leaf Display of the Effect Sizes at Posttesting on All Vocabulary Measures.

Stem	Overall Vocabulary [a]	Expressive Vocabulary [b]	Receptive Vocabulary [c]
1.0		3, 3	
.9	0, 1		
.8			
.7	3, 3	3	3
.6			7
.5	1, 3, 8	9, 9	8, 8
.4	4		1, 4
.3	9	2	3, 9
.2	6	3	7
.1	6	6, 8	3
.0			
-.0	5		
-.1	5, 9		5, 8, 9
-.2	6, 6		6
-.3			
-.4			
-.5			
-.6			
-.7			
-.8			8

Note. Combine the numbers under 'stem' with the numbers in the other columns to find all effect sizes for overall, expressive, and receptive vocabulary; for instance, negative effect sizes for 'overall vocabulary' were: -.05, -.15, -.19, -.26, and -.26. [a] Overall mean $d = .42$, $k = 16$ studies, $n = 626$ children ; [b] Overall mean $d = .59$, $k = 9$, $n = 322$; [c] Overall mean $d = .22$, $k = 15$, $n = 608$

When focusing on measures of expressive vocabulary in particular ($k = 9$; $n = 322$), Cohen's d was .59 ($SE = .08$; *95%* CI = .44, .75; $p < .001$), which is a moderate effect size. For the studies that reported receptive vocabulary measures ($k = 15$; $n = 608$), Cohen's d was small ($d = .22$, $SE = .09$; *95%* CI = .05, .39; $p < .01$). As the

95% CIs of these two outcome measures showed no overlap, the hypothesis that Dialogic Reading affects expressive vocabulary significantly more than receptive vocabulary was accepted. The number of missing studies that were needed to overturn these significant results were *N*fs = 92 for expressive and *N*fs = 9 for receptive vocabulary, respectively.

Explaining the Variability in Effect Sizes

A moderator analysis was conducted to test whether the intervention had a greater impact on younger children in preschool ($k = 10$, $n = 351$) than on older children in kindergarten ($k = 6$, $n = 275$). As is presented in Table 3.2, the overall vocabulary of preschool children benefited significantly more from the Dialogic Reading intervention ($Q = 7.14$, $p < .01$; $d = .50$, $SE = .12$; *95%* CI = .37, .64) than children in kindergarten ($d = .14$, $SE = .07$; *95%* CI = -.10, .37). This moderator could not be tested for expressive vocabulary outcomes, because the kindergarten subset consisted of less than four studies. With regard to receptive vocabulary, the moderator did not remain significantly different between subsets ($Q = 1.30$, $p > .05$).

Furthermore, we investigated whether children at risk benefited less from the Dialogic Reading intervention than children not at risk. Because SES was not clearly reported in all studies, risk status was based on the demographic variables of income or maternal education. Families designated as at risk ($k = 7$, $n = 208$) received governmental support (Lonigan & Whitehurst, 1998), had low incomes (Cronan, Cruz, Arriaga, & Sarkin, 1996), or had low educated mothers (Crain-Thoreson & Dale, 1999; Fielding-Barnsley & Purdie, 2002; 2003). Samples were designated as not at risk ($k = 9$, $n = 418$) if families had a modal income (Whitehurst et al., 1988), and mothers were educated at the tertiary (Arnold et al., 1994; Huebner, 2000b) or secondary level (Blom-Hoffman, O'Neill-Pirozzi, Volpe, Cutting, & Bissinger, 2006; Chow & McBride-Chang, 2003; Chow et al., 2007). The effect of Dialogic Reading did significantly differ between subsets ($Q = 9.52$, $p < .01$), with a moderate effect size for children not at risk ($d = .53$, $SE = .07$; *95%* CI = .40, .67) and only a small effect for children at risk ($d = .13$, $SE = .11$; *95%* CI = -.08, .35). When focusing solely on expressive vocabulary, the same significant differences in effect sizes were present ($Q = 6.80$, $p < .01$). When selecting receptive vocabulary outcomes, however, the moderator did not remain significant ($Q = .002$, $p > .05$).

Table 3.2
Meta-analytic Results of Dialogic Reading Intervention Studies and Moderators split for Overall vocabulary, Expressive and Receptive Vocabulary Outcome Measures.

	k	*n*	*d*	95% CI	*Q* [a]	*p*	I^2
Overall Vocabulary							56.01
Total set	16	626	.42	.16, .54	34.10	.00	
At risk					9.52	.00	
Yes	7	208	.13	-.08, .35	8.95	.18	32.94
No	9	418	.53	.40, .67	15.63	.05	48.83
Age group [b]					7.14	.01	
PreS	10	351	.50	.37, .64	18.60	.03	40.27
K	6	275	.14	-.10, .37	8.37	.14	51.60
Expressive Vocabulary							
Total set	9	322	.59	.44, .75	11.27	.19	29.01
At risk					6.80	.01	
Yes	4	96	.22	-.10, .54	0.13	.99	0.00
No	5	226	.71	.53, .89	4.34	.36	7.78
Age group					.66	.42	
Receptive Vocabulary							
Total set	15	608	.22	.05, .39	21.34	.09	34.40
At risk					1.14	.29	
Age group					1.30	.26	

Note. *k* = number of studies; *n* = total number of participants; *95%* CI = Confidence Interval; [a] *Q* statistic for moderator stands for effects of contrasts (*df* = number of subsets - 1); *Q* statistic for subset stands for homogeneity (*df* = k - 1); [b] *PreS* = Preschool; *K* = Kindergarten

Discussion

This meta-analysis tested the feasibility of an intervention designed to increase the quality of shared book reading among parents and their two- to six-year-old children. The meta-analysis demonstrates that enhancing the dialogue between parent and child during reading sessions strengthens the effects of book reading. However, the correlation between the intervention and a compound of linguistic skills was moderate ($r = .20$), explaining about 4% of the outcome measures in a set of 16 studies that included 626 children. When we restricted the analyses to studies that assessed expressive vocabulary ($k = 9$ studies, $n = 322$ children), the relation became stronger ($r = .29$), explaining about 8% of the variance. Apparently, not only does the exposure to a story promote language development, but it is also important that parents stimulate active involvement by eliciting verbal responses to the story with the help of open-ended questions. This outcome means it is likely that the quality of book reading is as important for language development as its frequency. The large effect size reported in Whitehurst et al.'s (1988) study was never replicated, but an authorial bias was not present as appears from the comparison of studies executed by Whitehurst and colleagues and studies carried out by other research groups as well as from a non-significant publication bias. Besides that, a cumulative analysis did not show a decreasing effect of the intervention with an increase in publication year; a common result of other meta-analyses in which pioneering studies often appear as outliers that are unlikely to be replicated (Bus & van IJzendoorn, 2004; Petticrew & Roberts, 2006).

Not all children need Dialogic Reading to profit from parent-child book sharing, however. We also found evidence supporting the hypothesis that Dialogic Reading with older children does not have as great an impact as Dialogic Reading with younger age groups. The older children in this set of studies, the five- to six-year-olds, scarcely benefited from Dialogic Reading (d = .14), explaining less than 1% of the variance. Insofar as Dialogic Reading causes additional effects, these effects were only manifested in the younger, two- to four-year old age group (d = .50). In the latter group, about 4% to 5% of the differences in outcome measures were explained by stimulating active child participation through Dialogic Reading. It is not plausible that a lack of challenge causes the low effect sizes in kindergarten children, but it is conceivable that parents fail to adapt the technique to older children. That is to say, a set of specific techniques has been developed for reading with children aged four to five years alongside a set for reading with children aged two to three years (Zevenbergen & Whitehurst, 2003). The set for older children targets more advanced skills by asking specific types of questions, evaluating and expanding on the child's responses, and having the child repeat the expanded phrases. However, the information provided in the articles was mostly not sufficient to decide whether researchers indeed adapted Dialogic Reading in order to implement more challenging interventions for the older children.

The finding that older and more experienced children may depend less on external support in understanding and enjoying the story is in line with the assumption that children internalize previous experiences and generalize those to new situations. As a result, kindergarten children need less help and support to remain attentive and to discover exciting parts of the stories, even when stories are new. It is possible that the same amount of dialogue takes place during parent-child interaction, but that older children depend less on adults because they are more inclined to initiate dialogue when they do not grasp the story content or are deeply affected by events. As is also suggested by the outcomes of studies comparing the amount of talk in older and younger age groups (Bus & van IJzendoorn, 1988; Martin, 1998), another option is that sessions become less dialogic with age. Older children may even prefer hearing the story without interruptions because they have sufficient linguistic skills and knowledge to sustain interest in the story without a parent focusing their attention by asking questions and providing explanations. In fact, they may not experience interruptions for questions as stimulating, but rather as annoying and interfering. This interpretation is further supported by the number of studies involving kindergarten children that reported negative results for Dialogic Reading, namely three out of six experiments, in contrast to two out of ten studies with negative results when participants were preschool children. Unfortunately, it is not possible to check this explanation because there are no studies demonstrating that older children indeed respond differently to Dialogic Reading compared with younger children.

Is Dialogic Reading a helpful technique to promote school readiness in children who are most in need of effective language promotion and pre-literacy experiences? In this context, any implementation of Dialogic Reading that includes parents with greater educational diversity should be considered (Huebner & Meltzoff, 2005). Parents with relatively low levels of education in particular could benefit from evocative, literacy-stimulating techniques, as dialogue during book sharing is not a self-evident phenomenon in low-educated families. A low-educated mother, in contrast to a better educated mother, often just explains details of the picture without any attempt to involve the child in thinking about the event (Arnold et al., 1994; Bus & van IJzendoorn, 1995; Cronan et al., 1996; Huebner & Meltzoff, 2005). It may therefore be hard for these parents to incorporate such behavior into book reading routines (e.g., Bus & van IJzendoorn, 1995; Heath, 1982; Ninio, 1980). For that reason, we tested whether effect sizes variedy as a function of risk status, a variable that, fortunately, was not confounded with age group in the present set of studies.

A unique result of this meta-analysis is that groups at risk for language and literacy impairments benefit less from Dialogic Reading than groups not at risk. Specifically, in the group at risk, we found a minimum effect size (explaining 1% of the variance), whereas in groups not at risk the effect size was substantial (7%). Two explanations for this disappointing result are germane. First, one could argue, in line with the so-called Matthew effect, that parents are required to have a strong educational background in order to use Dialogic Reading effectively (Fung et al., 2005). A critical test of this explanation is missing because there are no studies testing to what extent Dialogic Reading is actually realized in lower-and higher-educated families. Second, it may be that children at risk do not benefit from Dialogic Reading because making inferences (and similar requests) goes beyond their present abilities. Attempts to expand children's behavior may not be effective as long as the new behavior is not part of children's developing repertory of responses (Bus & de Jong, 2006). It is therefore imaginable that older children at risk benefit more from Dialogic Reading than younger children at risk but in view of the number of studies we were unable to test this hypothesis.

Cautions and Limitations

Because of a number of drawbacks, it seems important to replicate this meta-analysis a few years from now when the number of studies has grown. The present study included only a small set of studies and a moderate number of participants. Furthermore, to separate benefits of book reading and encouragement from the effects unique to the Dialogic Reading method, we need to be sure that the frequency of book reading to children is similar in the control and experimental groups and that the sessions in the experimental group are more interactive as a result of training parents in Dialogic Reading. Many studies lack control of what actually happens in the control and experimental groups. Data describing the behavior of the control group and/or intervention group are often missing or scant. Providing more descriptive data is advisable, even though a quality variable coding the amount of information relating to the content and frequency of book reading did not correlate significantly with outcome measures. Furthermore, we cannot exclude bias due to the fact that the program was voluntary in many studies. It is possible that parents attracted to the intervention were those who were more likely to carry out program requirements. Lastly, it was not always possible to be certain of the origin of the intervention effects. In some studies (e.g., Crain-Thoreson & Dale, 1999; Lonigan & Whitehurst, 1998) the home interventions were part of a larger study that also targeted teacher-child reading. Although we solely selected children in the separate home condition, it cannot be ruled out that the interventions in the school environment affected their outcomes.

Practical Implications

The literature suggests that Dialogic Reading has potential for enhancing the language development of very young children, thus increasing the readiness with which they enter school (e.g., Cutspec, 2004). The behavioral change in parents that occured while sharing books with children, coupled with the expressive language gains demonstrated by the children who participated in the studies, provides an early childhood intervention that is worth implementing in families. However, such a conclusion with far-reaching consequences for intervention programs is only partly supported by the results of this quantitative meta-analysis based on a review of 16 experiments. Our findings indicate that Dialogic Reading does not form a scaffolding of parent-child opportunities for early literacy development for all parents. For reasons to be addressed and specified in further research, this meta-analysis indicates that a book-reading intervention standardized on middle-class White or suburban samples may not be appropriate for lower class families, analogous to outcomes of experiments in libraries and other domains (e.g., Neuman & Celano, 2006). In particular, the present set of studies suggests that Dialogic Reading can change the home literacy activities of families with two- to four-year-old children but not those of families with children at greatest risk of school failure.

Appendix

Appendix 3.1

Characteristics of Studies Included in the Meta-Analysis.

First Author	*Year*	*Continent*	*At Risk*	*School*[a]	M_{age} *int. gr.*	*Design*[b]	*Training DR*[c]	*Duration (weeks)*	*Control gr.*[d]	*N* ($n_{exp.}$ & $n_{cont.}$)[e]	*Outcome*[f]	*ES (d)*[g]	*95% CI*
Arnold (*st 1*: DR vs C)[1]	1994	U.S.	No	PreS	28.20	E	Group	4	No int.	37 (23 & 14)	*RV*	.41	-.26, 1.08
											EV	.59	.11, 1.07
											OL	.53	.14, .92
Arnold (*st 2*: DR_Video vs C)[1]	1994	U.S.	No	PreS	29.70	E	Video	4	No int.	27 (14 & 13)	*RV*	.67	-.11, 1.45
											EV	1.03	.46, 1.59
											OL	.90	.44, 1.36
Blom-Hoffman	Unpub.	U.S.	No	PreS	44.50	E	Video	6	Books	18 (8 & 10)	*EV*	.73	-.23, 1.69
Chow (*st 1*: DR vs TR)[2]	2003	Asia	No	K	63.72	E	Indiv.	8	Books	41 (13 & 28)	*RV*	-.15	-.81, .51
Chow (*st 2*: DR vs C)[2]	2003	Asia	No	K	63.72	E	Indiv.	8	No int.	41 (14 & 27)	*RV*	-.19	-.84, .46
Chow (*st 1*: DR vs TR)[2]	Unpub.	Asia	No	K	63.11	E	Group	12	Books	55 (18 & 37)	*RV*	.44	-.13, 1.01
Chow (*st 2*: DR vs C)[2]	Unpub.	Asia	No	K	63.11	E	Group	12	No int.	55 (19 & 36)	*RV*	.39	-.17, .96
Crain-Thoreson	1999	U.S.	Yes	PreS	49.90	E	Group	8	No int.	19 (10 & 9)	*RV*	.13	-.78, 1.03
											EV	.18	-.46, .82
											OL	.16	-.36, .69
Cronan (*st 1*: 18 Instr. visits vs C)[3]	1996	U.S.	Yes	PreS	27.80	E	Indiv.	28	No int.	32 (21 & 11)	*RV*	.73	-.02, 1.48
Cronan (*st 2*: 3 Instr. visits vs C)[3]	1996	U.S.	Yes	PreS	28.00	E	Indiv.	28	No int.	31 (21 & 10)	*RV*	-.26	-1.01, .50

Fielding-Barnsley	2002	Australia	Yes	K	63.00	QE	Indiv.	8	No int.	34 (17 & 17)	*RV*	-.88	-1.59, -.18
											EV	.32	-.36, .10
											OL	-.26	-.74, .23
Fielding-Barnsley	2003	Australia	Yes	G1	70.20	QE	Indiv.	8	No int.	49 (26 & 23)	*RV*	.58	.00, 1.15
Huebner	2000b	U.S.	No	PreS	28.61	E	Group session	6	Books	115 (79 & 36)	*RV*	.27	-.13, .67
											EV	.59	.36, .83
											OL	.51	.31, .71
Lonigan (*st 1*: low compliance)	1998	U.S.	Yes	PreS	41.90	E	Indiv.	6	No int.	20 (9 & 11)	*RV*	.33	-.56, 1.22
											EV	.23	-.40, .85
											OL	.26	-.25, .77
Lonigan (*st 2*: high compliance)	1998	U.S.	Yes	PreS	46.90	E	Indiv.	6	No int.	23 (7 & 16)	*RV*	-.18	-1.07, .71
											EV	.16	-.46, .79
											OL	-.05	-.56, .47
Whitehurst	1988	U.S.	No	PreS	29.40	E	Indiv.	4	No int.	29 (14 & 15)	*RV*	.58	-.17, 1.32
											EV	1.03	.58, 1.48
											OL	.91	.53, 1.29

Note. [a] School type at the start of the study: preschool (PreS), kindergarten (K), first grade (G1). [b] Design: experimental (E), quasi-experimental (QE). [c] Type of Dialogic Reading parent training (video instruction, group session, individual training). [d] Control group: got no intervention (No int.) or received books but no dialogic reading instruction (Books). [e] *n*'s were adapted when studies included more than one intervention or control group and all groups met the inclusion criteria. [f] Outcome measures were indicators of *receptive vocabulary* (RV): Peabody Picture Vocabulary Test (PPVT, e.g. Dunn & Dunn, 1997), Bracken Basic Concept Scale (BBCS, Bracken, 1984); and indicators of *expressive vocabulary* (EV) that were combined into an overall EV-*d*: Expressive One-Word Vocabulary Test (EOWPVT, Gardner, 1981), expressive vocabulary subtest of the Illinois Test of Psycholinguistic Abilities (ITPA-VE, Kirk, McCarthy & Kirk, 1968), mean and/or total length of utterances by the child during reading sessions (MLU); *Oral Language (OL)*: effect sizes of indicators of receptive and expressive vocabulary were combined into an overall OL-*d*. [g] Based on posttest data: means and standard deviations. When posttest means and standard deviations were not reported, Cohen's *d* was calculated using: t-test (Cronan et al., 1996), covariate adjusted F-test and *p*-values (Arnold et al., 1994), and F-test statistic based on pre-post gain scores (Fielding et al., 2002). [1] The size of the control group was split into two groups in order to include results of both Dialogic Reading-interventions. [2] The size of the Dialogic Reading group was split into two groups in order to include both control groups. Means and standard deviations remained unchanged. [3] We excluded the one- to two-year-olds because outcome measures were based on parent reports. Test results of the three-year-olds were based on objective measures of receptive vocabulary and included in the meta analysis. The n's of the intervention and control group were estimated by adding two to the number of degrees of freedom of the reported t-test statistic before dividing the adjusted degrees of freedom by two. Then, the size of the control group was split into two groups in order to include results of both dialogic reading interventions.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

* Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology, 86,* 235-243.

Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129,* 195-215.

* Blom-Hoffman, J., O'Neill-Pirozzi, T., Volpe, R., Cutting, J., & Bissinger, E. (2006). Instructing parents to use dialogic reading strategies with preschool children: Impact of a video-based training program on caregiver reading behaviors and children's related verbalizations. *Journal of Applied School Psychology,* 23, 117-131.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2.* Englewood NJ: Biostat.

Bracken, B.A. (1984). *The Bracken Basic Concept Scale: Examiner's manual.* Columbus, OH: Bell & Howard.

Britto, P. R., Brooks-Gunn, J., & Griffin, T. M. (2006). Maternal reading and teaching patterns: Associations with school readiness in low-income African American families. *Reading Research Quarterly, 41,* 68-89.

Bus, A. G., & de Jong, M. T. (2006). Relationships in book sharing: A developmentally appropriate way to foster pre-academic growth. In S. Rosenkoetter & J. Knapp-Philo (Eds.), *To read the world: Language and literacy in the first 3 years* (pp. 123-144). Washington, DC: Zero-to-Three Press.

Bus, A. G., Leseman, P. P., & Keultjes, P. (2000). Joint book reading across cultures: A comparison of Surinamese-Dutch, Turkish-Dutch, and Dutch parent-child dyads. *Journal of Literacy Research, 32,* 53-76.

Bus, A. G., & van IJzendoorn, M. H. (1988). Mother-child interactions, attachment, and emergent literacy: A cross-sectional study. *Child Development, 59,* 1262-1272.

Bus, A. G., & van IJzendoorn, M. H. (1995). Mothers reading to their three-year-olds: The role of mother-child attachment security in becoming literate. *Reading Research Quarterly, 30*, 998-1015.

Bus, A. G., & van IJzendoorn, M. H. (2004). Meta-analysis in reading research. In: N. Duke & M. Mallette (eds.), *Literacy research methods* (pp. 227-252). New York: Guilford Press.

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1-21.

* Chow, B. W.-Y., & McBride-Chang, C. (2003). Promoting language and literacy development through parent-child reading in Hong Kong preschoolers. *Early Education and Development, 14,* 233-248.

* Chow, B. W.-Y., McBride-Chang, C., Cheung, H., & Chow, C. S.-L. (2007). Dialogic reading and morphology training in Chinese children: Effects on language and literacy. Manuscript submitted for publication.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Cooper, H., & Hedges, L. V. (Eds) (1994). *The handbook of research synthesis.* New York: Russell Sage Foundation.

* Crain-Thoreson, C., & Dale, P. S. (1999). Enhancing linguistic performance: Parents and teachers as book reading partners for children with language delays. *Topics in Early Childhood Special Education, 19,* 28-39.

* Cronan, T. A., Cruz, S. G., Arriaga, R. I., & Sarkin, A. J. (1996). The effects of a community-based literacy program on young children's language and conceptual development. *American Journal of Community Psychology, 24,* 251-272.

Cutspec, P. A. (2004). Influences of dialogic reading on the language development of toddlers. *Bridges; Practice-Based Research Syntheses, 2,* 1-12. Retrieved February 21, 2007.

Dale, P. S., Crain-Thoreson, C., Notari-Syverson, A., & Cole, K. (1996). Parent-child book reading as an intervention technique for young children with language delays. *Topics in Early Childhood Special Education, 16,* 213-235.

DeTemple, J. & Snow, C. E. (2003). Learning words from books. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 16-36). Mahwah, NJ: Erlbaum.

Dunn, L. & Dunn, L. (1997). *Peabody Picture Vocabulary Test* (PPVT-III, 3rd ed.). Circle Pines, MN: American Guidance Service.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method for accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89-98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455-463.

* Fielding-Barnsley, R., & Purdie, N. (2002). Developing pre-literacy skills via shared book reading: The assessment of a family intervention program for preschool children at risk of becoming reading disabled. *Australian Journal of Learning Disabilities, 7,* 13-19.

* Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure. *Support for Learning, 18,* 77-82.

Fung, P. C., Chow, B. W. Y., & McBride-Chang, C. (2005). The impact of a dialogic reading program on deaf and hard-of-hearing kindergarten and early primary school-aged students in Hong Kong. *Journal of Deaf Studies and Deaf Education, 10,* 82-95.

Gardner, M. F. (1981). *Expressive One-Word Picture Vocabulary Test*. Novata, CA: Academic Therapy Publications.

Gormley, S., & Ruhl, K. L. (2005). Dialogic shared storybook reading: An instructional technique for use with young students in inclusive settings. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 21,* 307-313.

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15,* 75-90.

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11,* 49-76.

Huebner, C. E. (2000a). Community-based support for preschool readiness among children in poverty. *Journal of Education for Students Placed at Risk, 5*, 291-314.

* Huebner, C. E. (2000b). Promoting toddlers' language development through community-based intervention. *Journal of Applied Developmental Psychology, 21*, 513-535.

Huebner, C. E., & Meltzoff, A. N. (2005). Intervention to change parent-child reading style: A comparison of instructional methods. *Journal of Applied Developmental Psychology, 26,* 296-313.

Juel, C. (2006). The impact of early school experiences on initial reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol.2, pp. 410-426). New York: Guilford Press.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). *Illinois Test of Psycholinguistic Abilities.* Urbana: University of Illinois Press.

Laakso, M. L., Poikkeus, A. M., &Lyytinen, H. (1999). Shared reading interaction in families with and without genetic risk for dyslexia: Implications for toddlers' language development. *Infant and Child Development, 8,* 179-195.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis.* Thousand Oaks, CA: Sage.

Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention, 22,* 306-322.

* Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention of preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13,* 263-290.

Martin, L. E. (1998). Early book reading: How mothers deviate from printed text for young children. *Reading Research and Instruction, 37,* 137-160.

Neuman, S. B., & Celano, D. (2006). The knowledge gap: Implications of leveling the playing field for low-income and middle-income children. *Reading Research Quarterly, 41,* 176-201.

Ninio, A. (1980). Picture-book reading in mother-infant dyads belonging to two subgroups in Israel. *Child Development, 51,* 587-590.

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language, 5,* 1-15.

Park, J. Y. K. (2006). *Facilitating young children's second language development through picture book sharing*. Unpublished doctoral dissertation, University of Kansas, USA.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences; A practical guide.* Oxford, England: Blackwell Publishing.

Raikes, H., Pan, B. A., Luze, G., Tamis-LeMonde, C. S., Brooks-Gunn, J., Banks Tarullo, J.C., Raikes, H. A., & Rodriguez, E. T. (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development, 6,* 193-235.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183-192.

Silvén, M., Ahtola, A., & Niemi, P. (2002). Early words, multiword utterances and maternal reading strategies as predictors of mastering word inflections in Finnish. *Journal of Child Language, 30,* 253-279.

Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. *Reading Research Quarterly, 20,* 458-481.

Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading: A systematic extension to Mexican day care. *Developmental Psychology, 28,* 1106-1114.

Verhallen, M. J. A. J., Bus, A. G., & De Jong, M. T. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology, 98,* 410-419.

Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology, 93*(2), 243–250.

Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30,* 679-689.

Whitehurst, G. J., Epstein, J. N., Angel, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology, 86,* 542-555.

* Whitehurst, G. J., Falco, F. L., Lonigan, C. J., & Fischel, J. E., DeBaryshe, F. D., Valdez-Menchaca, M. C., et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24,* 552-559.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69, 848-872.*

Whitehurst, G. J., Zevenbergen, A. A., Crone, D. A., Schultz, M. D., Velting, O. N., & Fischel, J. E. (1999). Outcomes of an emergent literacy intervention from Head Start through second grade. *Journal of Educational Psychology, 91,* 261-272.

Wilson, D.B. (2001). Effect size determination program. Retrieved February 28, 2007.

Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic reading: A shared picture book reading intervention for preschoolers. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 177-200). Mahwah, NJ: Erlbaum.

4

Interactive Book Reading in Early Education:

A Tool to Stimulate Print Knowledge as well as Oral Language

Abstract

This meta-analysis examines to what extent interactive storybook reading stimulates two pillars of learning to read: vocabulary and print knowledge. We quantitatively reviewed 31 (quasi)experiments (*N* = 2,049 children) in which educators were trained to encourage children to be actively involved before, during, and after joint book reading. A moderate effect size of *d*=.54 (CI = .33, .74) was found for oral language skills, implying that both quality and quantity of book reading in classrooms are important. Although teaching print-related skills is not part of interactive reading programs, 7% of the variance in kindergarten children's alphabetic knowledge could be attributed to the intervention. The study also shows that findings with experimenters were simply not replicable in a natural classroom setting. Further research is needed to disentangle the processes that explain the effects of interactive reading on children's print knowledge and the implementation strategies that may help transferring intervention effects from researchers to children's own teachers.

Based on:
Mol, S. E., Bus, A. G., & De Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research, 79,* 979-1007. Doi: 10.3102/0034654309332561

Introduction

Exposure to books is considered to be a major source for developing one of the pillars of learning to read: vocabulary (e.g., Bus, Van IJzendoorn, & Pellegrini, 1995; Juel, 2006; Scarborough & Dobrich, 1994; Sénéchal & LeFevre, 2002). The nature of a text, the quality of the reading style, and the number of times a book is reread seem to be important contributors to young children's vocabulary development (e.g., Dickinson & Smith, 1994; McKeown & Beck, 2006; Reese & Cox, 1999). It is under discussion to what extent book reading also fosters the second pillar, print knowledge. There is ample evidence for the hypothesis that children spontaneously ignore the print during storybook reading (Bus & Van IJzendoorn, 1988; Evans & Saint-Aubin, 2005; Justice, Skibbe, Canning, & Lankford, 2005; Yaden, Smolkin, & Conlon, 1989). On the other hand, as children grow older and become more proficient in print knowledge, they may feel more attracted to letters and sounds in books (e.g., Mason, 1992; Morris, Bloodgood, Lomax, & Perney, 2003; Roy-Charland, Saint-Aubin, & Evans, 2007). In other words, even though evocative techniques, informative feedback, and sensitivity to a child's abilities include incentives mostly for children's oral language skills and not for print knowledge (e.g., Justice & Ezell, 2002; Whitehurst, Epstein, et al., 1994), the story text itself might trigger children's attention to the print, perhaps resulting in interactions with grown-ups that go beyond story understanding.

A recent meta-analysis of the effects of interactive book reading experiments in the family showed that 4% of the variance in vocabulary growth was explained by the additional effects of Whitehurst's (Whitehurst, Falco, Lonigan, & Fischel, 1988) "*Dialogic Reading*"- technique (Mol, Bus, De Jong, & Smeets, 2008). We hypothesized that Dialogic Reading provides a venue for focused language exchange, enabling responses to children's utterances and thinking processes as well as exposure to more formal adult language (e.g., Raikes et al., 2006). The data provided by this set of studies did not enable us to test the extent to which print knowledge was affected by the intervention, however, as there were hardly any studies that included measures such as alphabet knowledge, phonological sensitivity, or orthographic awareness. Strikingly, two subgroups did not appear to benefit from the intervention: The oral language skills of kindergarten children as well as children at risk for language and literacy impairments benefited less from interactive parent-child book reading (Mol et al., 2008). Because dialogue during shared reading is hardly observed in families at risk (e.g., Bus & Van IJzendoorn, 1995; Heath, 1982; Ninio, 1980), low-educated parents might have experienced difficulty with incorporating the trained techniques. On the other hand, expectations and methods may be pitched too low for older children. Too much talking might have a depressing effect on learning in more advanced groups. As teachers appear to provide more cognitively demanding talk about books than parents, the literacy environment at school might be more stimulating for these groups of children (Dickinson & Smith, 1994). Covering book reading research in school settings until 1995, Blok (1999) tested the effects of book reading frequency on language and reading development in 2.5-to 7.5-year-old children, but the studies he included in his meta-analysis did not provide sufficient information about the quality of book reading such as the reading style of the teachers. Therefore, the current meta-analysis elaborates on the gaps in these previous meta-analyses by exploring to what extent Dialogic Reading – taking the form of (a) the use of evocative techniques that encourage the child to talk about pictured materials; (b) informative feedback that highlights the differences between what the child has said and what he or she might have said; and (c) an adaptive adult who is sensitive to the child's developing abilities (e.g., Whitehurst et al., 1988) – before, during, and after reading storybooks affects children's language acquisition as well as print knowledge.

We selected studies in which teachers and/or graduate students were instructed to implement an interactive reading intervention in preschool or kindergarten classrooms. Insofar as the interventions did not use Whitehurst and colleagues' (1988) techniques of Dialogic Reading, teachers and experimenters were trained in applying similar reading techniques: to prompt child responses by asking open-ended questions or making comments, and to support children's enthusiasm and learning opportunities by providing positive reinforcement or relating the story text to their real life experiences. Mostly, teachers or experimenters received handouts that summarized the learned techniques as well as (suggestions for) storybooks. Alternatively, scripted questions or comments were added to storybooks in order to promote the use of similar interactive prompts and responses in each classroom (Brabham & Lynch-Brown, 2002; Kertoy, 1994; Mautte, 1990; Van Kleeck, Vander Woude, & Hammett, 2006). Some teachers were repeatedly observed in the classroom and coached by the researcher to ensure that they applied the reading strategies in various situations (Aram, 2006; Droop, Peters, Aarnoutse, & Verhoeven, 2005; Wasik & Bond, 2001; Wasik, Bond, & Hindman, 2006). In several studies, interactive book reading was accompanied by kits with materials that focused on book-related vocabulary, games with rhyme or letters, and painting or dramatizing the stories (Aram, 2006; Aram & Biron, 2004; Droop et al., 2005; Karweit, 1989; Wasik & Bond, 2001; Wasik et al., 2006). As these additional activities were integrated in the classroom environment, they were expected to foster vocabulary and print knowledge beyond the interactive reading sessions. For children at risk in particular, repeatedly interacting with storybooks might be an important extra stimulant. That is, we do not expect that reading in the classroom can add to the rich home literacy environment that children who are not at risk are likely to experience (e.g., Adams, 1990; Hart & Risley, 2003).

We anticipated greater gains for experiments in which experimenters read to the children than for interventions that were executed by the children's own teachers. Compared to researchers who are well informed about literacy acquisition, teachers may be less familiar with ways to promote literacy and with theories behind interventions. Besides, teachers may have less time and energy to invest in a program that has to be combined with everyday responsibilities (Aram, 2006). It is of great significance, however, that intervention effects induced by researchers can be transferred or generalized to classroom conditions (e.g., Fuchs et al., 2001; Shernoff & Kratochwill, 2007). It seems that teachers are more likely to implement innovations when the programs are well specified, include attractive and user-friendly materials, and are accompanied with training and technical assistance such as coaching or personalized consultation prior to and during the implementation phase (Rohrbach, Grana, Sussman, & Valente, 2006).

Group size may be another important moderator. That is, it can be questioned whether it is possible to engage all participants in group conversations that are challenging as well as comprehensible to children (Dickinson & Sprague, 2001). Sessions involving the entire class require a level of attention that at-risk youngsters are more likely to lack due to fewer opportunities to practice focused attention in other settings (Bodrova & Leong, 2006; Diamond, Barnett, Thomas, & Munro, 2007). On the other hand, as group sessions offer ample opportunity to observe and interact with more literate peers, we expected that the oral language skills of children at risk may improve from shared reading in (small) group sessions. Morrow and Smith (1990) showed that reading to children in small groups offered as much interaction as one-to-one reading, and led even to greater gains in story comprehension than individual sessions. Teachers seem to provide more positive comments and spend more time redirecting the discussion to the story when they are reading to small groups (Karweit & Wasik, 1996). In addition, children's receptive vocabulary is especially thought to improve as a result of repeatedly reading the same storybook because of the additional opportunities to encode, associate, and store novel information due to several exposures (Biemiller & Boote, 2006; Moschovaki & Meadows, 2004; Nielsen, 1993; Sénéchal, 1997).

In sum, by quantitatively and systematically summarizing (quasi-)experiments that examined the effects of interactive reading in educational settings, we addressed the following research questions:

1) Does trained interactive teacher behavior as a part of book reading improve young children's language and print-related skills, or does this behavior not add anything to the effects of joint book reading? We expected that children in the experimental groups would learn more than control-group children who were read to without a special focus on interaction.
2) Are effect sizes of interactive reading as great for print knowledge as oral language? We expected oral language skills to show greater gains than print-related skills in younger and hence less proficient children, whereas we hypothesized that print knowledge would be affected more in kindergartners.

3) Which conditions benefit the efficacy of an interactive reading intervention in the classroom? First, are interventions carried out by experimenters more effective than those implemented by teachers? Second, is reading in small groups more effective than whole-group reading or individual sessions? Third, is there support for the assumption that extra opportunities to use book vocabulary during play, art, or drama activities add to the effects of book reading, as Karweit and Wasik (1996) suggest? Fourth, are at-risk groups especially susceptible to interactive reading interventions, taking into account that they receive fewer incentives at home (Raikes et al., 2006)?

Possible methodological confounders, such as publication status, year of publication, design, and experiment fidelity, were examined as well.

Method

Search Strategy and Selection Criteria

This meta-analysis examines the effect of interactive book reading on the oral language and print knowledge of children not yet reading conventionally. To obtain eligible studies, social science research databases (PsycINFO, ERIC, Dissertation Abstracts International, WebSPIRS, C2-SPECTR, and the Best Evidence Encyclopedia) were searched up to December 2007, using different combinations of the keywords: *(dialogic/interactive) read*, intervention/program, teacher, classroom, early education, daycare, preschool,* and/or *kindergarten,* with *vocabulary, language acquisition/growth, story comprehension, (early/emergent) literacy, (print/alphabet*/letter) knowledge,* and *phon* / orthograph* awareness/ sensitivity* as dependent variables. We also used the so-called snowball method by identifying eligible studies within the references of the collected articles.

Studies were included when they met the following criteria: (a) the study used an interactive, shared reading intervention with open-ended questions, prompts, comments, and positive reinforcement in encouraging children to become actively involved before, during, and after storybook reading; (b) the program was implemented in day care centers, preschool, kindergarten, or first-grade classrooms, and was not part of a larger intervention that specifically targeted the teaching of literacy concepts such as phonological sensitivity or orthographic awareness; (c) teachers, teacher aides, and/or research assistants were trained in using interactive reading techniques with individual or groups of children; (d) participants had no mental, physical, or sensory handicaps and were pre-conventional readers; (e) outcome variables included at least one objective measure of vocabulary or story comprehension; (f) a (quasi-)experimental design was applied, randomly assigning children to either an experimental or control group on individual, school, or classroom level; (g) children in the control group attended the regular school program, not including interactive reading; and (h) articles were published or unpublished, as far as the language of the article could be interpreted and a sufficient amount of statistical information was reported to determine effect sizes. Combined home and school interventions were eligible when separate data for experimental children in a single teacher group were presented.

Studies were excluded when the reading sessions were not the main focus, but book reading motivated teaching vocabulary (e.g., Ard & Beverly, 2004; Beck & McKeown, 2007; Biemiller & Boote, 2006; Collins, 2005; Elley, 1989; Justice, Meier, & Walpole, 2005; Penno, Wilkinson, & Moore, 2002; Sénéchal, Thomas, & Monker, 1995) or reading strategies and print concepts (e.g., Justice & Ezell, 2002; McCormick & Mason, 1986; L. M. Phillips, Norris, Mason, & Kerr, 1990). Furthermore, some studies could not be included because relevant data for separate interventions were not reported (e.g., Hargrave & Sénéchal, 2000; Whitehurst, Epstein, et al., 1994; Whitehurst, Zevenbergen et al., 1999) or because no control group was included (e.g., Morrow & Smith, 1990; Reese & Cox, 1999; Van Elsäcker & Verhoeven, 1997). The study of Brabham and Lynch-Brown (2002) was only partly included. That is, we excluded data regarding third graders and an experimental group in which children were not allowed to interact during the storybook sharing. Of the two intervention groups in Kertoy's study (1994), we included only the experimental group in which adults asked questions because these techniques reflected our inclusion criteria best. Because Mautte (1990) presented composite scores for print knowledge outcomes instead of separate scores, we included only her oral language measure.

Coding Process

As study or methodological characteristics, we coded publication year, publication status (1. published, 2. unpublished), sample size, design (1. experiment, 2. quasi-experiment[1]), and experiment fidelity (check up on reading techniques and frequency, in experimental and control groups). To test which populations benefited most from the intervention, we coded the language of the shared reading sessions (1. English, 2. other), school type (1. preschool, 2. kindergarten), and risk status (1. at risk for language and literacy impairments, 2. not at risk) of the participating children. Intervention characteristics were coded as another group of moderators, among which were the following: characteristics of the adult who carried out the intervention (1. teacher, 2. experimenter), size of the groups in which book reading took place (1. individual, 2. small group [max. 5 children], 3. large group), type of intervention program (1. Dialogic Reading in accordance with Whitehurst et al. (1988), 2. interactive reading without extra activities, 3. interactive reading with extra book-related classroom activities), information

[1] A study was coded as an experiment when each individual child was randomly assigned to a control or experimental group. Studies that randomized on classroom- or school-level but reported results on a subject level were treated as quasi-experiments.

about activities in the control group (1. intervention, such as reading the same books as the experimental group without interaction, 2. no intervention (such as stimulating play or the standard curriculum), and the duration of the intervention (in weeks, and the number of, recommended and/or mean, interactive reading sessions).

To calculate Cohen's *d* effect sizes, we gave preference to computing those with the help of posttest means and standard deviations for oral language measures (receptive vocabulary, expressive vocabulary, story comprehension, and syntax) and print-related skills (alphabetic knowledge, phonological sensitivity and orthographic awareness). Because some studies presented only *F*-values (Wasik & Bond, 2001), means corrected for pretest scores (Morrow, 1988; Morrow, 1989), or gain scores (Aram, 2006) instead of "raw" posttest data, an extra moderator was created to test the effect of positively biased outcomes. Analyses showed that the mean effect sizes slightly decreased when studies with the adjusted scores were excluded. However, in broad lines, the outcomes were similar to those found for the whole sample. We decided therefore to estimate mean effect sizes with all included studies.

When authors reported two independent experiments within one article (Aram, 2006; Droop et al., 2005; Karweit, 1989; Lonigan & Whitehurst, 1998), both studies were treated and coded separately[2]. When studies presented two parallel comparisons with one control group and all groups met the inclusion criteria, we split the study and adapted the sample sizes without adjusting the outcome values. For example, Lamb (1986), Lonigan, Anthony, Bloomfield, Dyer, and Samwel (1999), and Mautte (1990) included one experimental and two control conditions: a control group in which children were read to from the same books as the experimental group without interaction, next to a control group that attended solely the standard preschool curriculum between the pre- and posttest. To compare the effect of the intervention with both control groups, we divided the sample size of the experimental group by two and treated outcomes of the comparison with the control groups as two separate studies. On the other hand, Morrow (1988) was interested in the effect of repeated versus onetime book reading, so she included two experimental groups and one control group. Therefore, we split the sample size of the control group and included all the children who were read to interactively. In all four studies in which the sample sizes were adjusted, the means and standard deviations remained unchanged.

[2] Droop, Peters, Aarnoutse, and Verhoeven (2005) reported two independent experiments in which they trained teachers to read interactively to children classified as at risk and not at risk. Although all participating children were read to in large groups, the authors treated children at risk and not at risk as separate groups in their analyses – without reporting overall means and standard deviations for the control and experimental groups. We included all five comparisons as independent studies as analyses in which we excluded these samples showed that mean effect sizes were not affected more than could be expected by a decrease in power.

As a control, we excluded one of the four samples of each study that we split to analyze only independent studies with original sample sizes. The main effects did not differ from the outcomes that included the complete set of studies.

Two independent coders both coded all derived studies. We agreed completely on the studies that did not meet the inclusion criteria and had to be excluded. Of the 31 included studies, the agreement across study characteristics and moderators ranged from .67 to 1.00, resulting in a *mean* κ of .93 (M = 98%). The experiment fidelity scale that we calculated consisted of a sum of four items: whether the researchers checked (by means of self-reports or audiotaped reading sessions) (a) the use of trained interactive reading techniques in the experimental group, (b) the quality of reading sessions within the control group, and (c) the frequency of book reading in the experimental and (d) in the control group. Because not all studies reported this information clearly, this scale was more difficult to code reliably but the level of agreement was still satisfactory (M = 85%; κ = .71, *range* = .59 – .86). All discrepancies between coders were discussed and corrected. Authors were contacted when some uncertainties could not be clarified after carefully reading their article.

Statistical Analysis

We quantified the added value of interactive reading on young children's language and literacy development by using the Comprehensive Meta-Analysis program (Version 2.2; Borenstein, Hedges, Higgins, & Rothstein, 2005). To give greater weight to studies with larger sample sizes, the standardized differences between the means were determined by using weights based on the inverse of the variance (Lipsey & Wilson, 2001). Per study, effect sizes were first computed for receptive and expressive vocabulary outcomes separately. We combined these aggregated vocabulary scores with syntax and story comprehension measures to estimate a composite effect size for oral language. For print knowledge, we calculated separate effect sizes for alphabetic knowledge (e.g., tests measuring concepts of print, letter names), phonological sensitivity (e.g., rhyme, alliteration, blending, elision), and orthographic awareness (e.g., name/word writing). A positive sign in the Cohen's d column of Appendix 4.1 indicates a favorable outcome for the intervention program, with a d of .20 interpreted as a small, .50 as a moderate and .80 as a large effect size (Cohen, 1992). Put differently, a d of .50 indicates that the interactive reading has moved a child to the 69th percentile on average, compared to a child in the control group (Bus & Van IJzendoorn, 2004). The binominal effect size display, computed from the formula .50 ± (r / 2), indicates to what extent the prediction of children's language and literacy outcomes is enhanced by the intervention.

Overall effect sizes and 95% confidence intervals (CIs) around the point estimates were based on random effects models because such an approach is a conservative solution to deal with heterogeneity (Ioannidis, Patsopoulos,

& Evangelou, 2007; Viechtbauer, 2007). In such a model it is assumed that the variability beyond subject-level sampling error is derived from random differences among studies whose sources cannot be identified (Lipsey & Wilson, 2001). Therefore, another random component (reflected by τ) is included in addition to subject-level sampling error, resulting in wider CIs. For moderator analyses, the random effects model was applied as well. Contrasts of methodological, population, and intervention characteristics were tested and presented only when all cells within the subset contained at least four studies (Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003). A significantly different effect size in study outcome was determined by a significant $Q_{between}(df)$-value.

Because studies with non-significant findings or small sample sizes are less likely to be published and we have not located any unpublished reports except for dissertations, we examined this potential publication bias graphically by funnel plot analysis. We also calculated fail-safe numbers (*N*fs), indicating the number of (unpublished) studies that are needed to overturn a significant result (Lipsey & Wilson, 2001).

Preliminary Testing of Potential Biases

For all six effect size composites, no outliers were identified on the basis of standardized *z*-values larger than 3.26 or smaller than -3.26 ($p < .001$), neither did we find evidence for any methodological biases. That is, publication status did not bias the effects of the intervention ($k_{dissertation} = 5$, $k_{published\ study} = 26$; $Q(1) = .21$, $p > .05$). The effect sizes for all language and print knowledge outcomes did not significantly differ for true experiments ($k = 19$) and quasi-experimental studies ($k = 12$; $Q(1) = 1.69$, $p > .05$). The fidelity scale that we developed to examine whether more controlled experiments would reveal stronger results than studies that hardly checked the content and/or frequency of book reading sessions in the experimental and control conditions revealed a statistically significant difference between quasi-experimental and experimental designs ($t(30) = 3.37$, $p < .01$). That is, experiments received significantly higher experiment fidelity scores than quasi experiments. When we entered all sum scores of the experiment fidelity scale ($M = 3.81$; $SD = 2.63$; *range* = 1 – 8) into a meta-regression analysis (unrestricted maximum likelihood), there was no evidence for significant regression models, implying that children's outcomes are not affected by the experiment fidelity. Due to inadequate descriptive data, we were unable to opt for a more elegant solution by accounting for the nonindependence of observations within schools or classes (Hedges, 2007). Furthermore, a cumulative analysis did not reveal a decreasing effect of the intervention with an increase in publication year.

Results

The final set of intervention studies targeting interactive reading in educational settings comprised 31 studies. Sixteen studies tested at least one print-related skill next to an oral language outcome. In sum, 2,025 children ($N_{Experimental\ Group}$ = 1,016; $N_{Control\ Group}$ = 1,009) were studied, with a mean sample size of 65 children (*SD* = 56.10, *range* = 13 – 248). Specifically, 1,030 participants attended day care or preschool programs, and 995 were in kindergarten, of which 1,501 were read to by their teachers and 524 by experimenters. Insofar as articles provided school details (less than half), it appeared that children attended the educational setting at least half a day. Children were exposed to an average of 42.3 interactive reading sessions (*SD* = 33.31, *range* = 4 – 66). For specific study characteristics and unweighted effect sizes per outcome, see Appendix 4.1.

In the first subsection, we examined the additional effects of interactive reading on oral language and print knowledge. Second, we tried to explain the variability in effect sizes on the basis of intervention characteristics. As regards the participants, it should be noted in advance that 27 out of 31 studies targeted students classified as at risk (*n* = 1,515), including children qualified for public subsidy of day care costs; attending Head Start, Title I, or similarly funded classrooms; and/or scoring below national norms on early literacy tasks. When we left out the four studies with children not at risk, mean effect sizes were similar to the analyses that comprised all samples (see Table 4.1).

Table 4.1
Mean Effect Sizes of All Six Outcome Measures in the Overall and the At-Risk Sample

	Overall Sample				*At-Risk Sample*			
	k	*d*	95% CI	*N*fs	*k*	*d*	95% CI	*N*fs
Oral Language	31	0.54***	.33, .74	1724	27	0.57***	.36, .78	1448
Expressive Vocabulary	20	0.62***	.29, .95	503	17	0.72***	.33, 1.10	518
Receptive Vocabulary	23	0.45***	.22, .68	463	20	0.48***	.12, .67	335
Print Knowledge								
Alphabet Knowledge	13	0.39**	.16, .62	112	11	0.40**	.12, .67	61
Phonological Sensitivity	13	0.43***	.25, .62	332	11	0.52***	.29, .76	246
Orthographic Awareness	9	0.41***	.20, .62	83	7	0.36**	.10, .62	27

** $p < .01$, ***$p < .001$

Oral Language and Print Knowledge Outcomes

To examine to what extent children's oral language would benefit more from an interactive reading intervention than print-related skills, separate meta-analyses on each outcome were conducted.

The interactive reading intervention had a moderate effect on the oral language skills, explaining 6% of the variance ($d = .54$, $p < .001$; *95%* CI = .33, .74). Expressive vocabulary was especially affected by interactive reading ($k = 20$, $n = 1{,}350$; $d = .62$, $p < .001$; *95%* CI = .29, .95). However, this effect size did not differ significantly from receptive vocabulary because the 95% CIs showed considerable overlap ($k = 23$; $n = 1{,}765$; $d = .45$, $p < .001$; *95%* CI = .22, 68). To overturn these results into non-significant effect sizes a substantial number of missing studies have to be located or executed: For oral language outcomes the fail-safe number equaled 1,724; for expressive vocabulary, 503; and for receptive vocabulary, 463, respectively. In Figure 4.1, the effect sizes for all studies are displayed graphically.

Print knowledge was split into three subcategories: alphabetic knowledge ($k = 13$, $n = 1{,}170$), phonological sensitivity ($k = 13$, $n = 1{,}105$), and orthographic awareness ($k = 9$; $n = 880$). The additional effects of interactive reading on these skills explained about 4% to 5% of the variance and can be interpreted as modest, varying from $d = .39$ for alphabetic knowledge ($p < .01$; *95%* CI = .16, .62; *N*fs = 112) to $d = .43$ for phonological sensitivity ($p < .001$; *95%* CI = .25, .62; *N*fs = 332) and $d = .41$ for orthographic awareness ($p < .001$; *95%* CI = .20, .62; *N*fs = 83).

When the precision (1/SE) was plotted against the standardized difference in means, symmetry around the point estimate appeared to be present for oral language, expressive and receptive vocabulary, alphabetic knowledge, and orthographic awareness. However, a publication bias was detected in the plot reflecting phonological sensitivity outcomes. That is, four studies in the bottom left-hand corner had to be added in order to find symmetry around the point estimate. We used the trim-and-fill method to calculate the effect of this potential data censoring: Trimmed studies were replaced, and their missing counterparts were imputed as mirror images of the trimmed outcomes (Duval & Tweedie, 2000a, 2000b). Adding four studies that appeared to be missing in order to obtain symmetry resulted in an adjusted effect size of $d = .25$ (*95%* CI = .06, .45), suggesting the effect size to be closer to .25 than .43.

In sum, we can accept our first hypothesis: Children's language and print-related skills seem to improve as a result of interactive reading interventions. Second, we wondered whether all skills would be affected equally. Because the 95% CIs of the effect sizes of all outcome measures showed overlap, we had to reject our hypothesis that oral language skills – and expressive vocabulary in particular – would gain the most from the interaction. Although adults were not instructed to comment on letters, phonemes, or writing concepts during storybook reading, the results suggest that being read to interactively can be seen as an incentive for improving both oral language and print knowledge.

Specifically, we expected that the intervention would not affect all children's print-related skills to the same extent. That is, as children grow older, they might spontaneously pay more attention to print. To test this hypothesis, we compared older kindergarten groups with younger preschool groups. Twenty-one studies

Figure 4.1

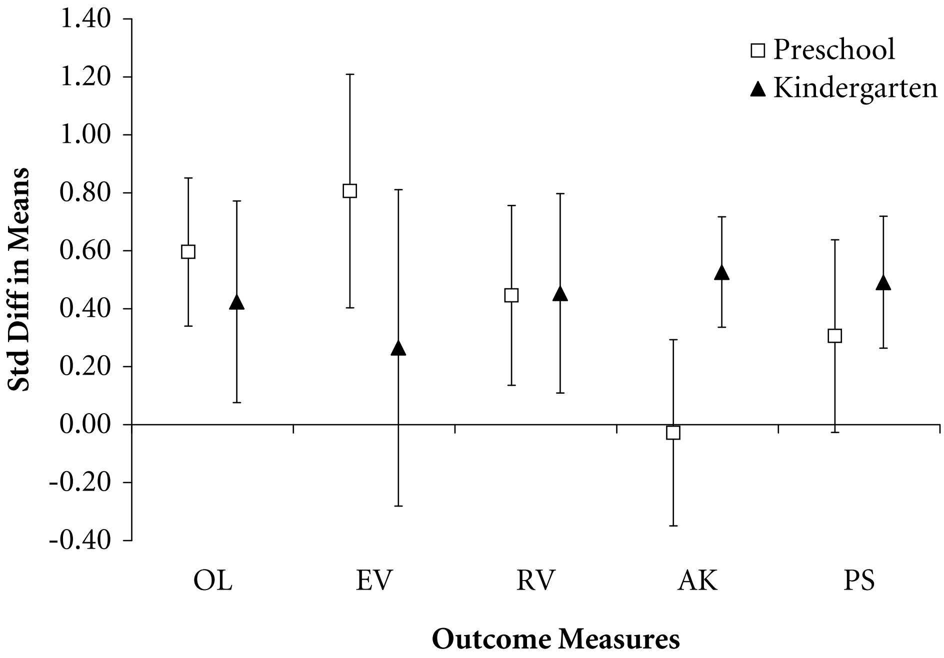
Stem-and-Leaf Display of the Effect Sizes per Study at the Posttest on all Outcome Measures

Stem	OL	EV	RV	AK	PS	OA
2.7	1	1				
2.6						
2.5						
2.4						
2.3						
2.2	7	7				
2.1						
2.0		6				
1.9						
1.8						
1.7		5				
1.6						
1.5	7		2			
1.4	6		1		8,8	
1.3	5					
1.2			1			
1.1	8					
1.0	2	7	8		4	
0.9	3	3	2	2		
0.8	4					5
0.7		1	8	3		1
0.6	1,6,9	0	2	0,0,1,2,4,6	6,7	
0.5	3,4	4	2,3,4		3	2,9
0.4	6,7,7	5	6,7		7	
0.3	4,5	4			7	5,9
0.2	0,5	4	4	5	8,9	9
0.1	0,9	5	5,6,6		4	
0.0	4	4,6,9	3,6	7,8	2	
-0.0	8	7	1,4			
-0.1	3,3,3	6,8	0,9			0
-0.2	1			6		1
-0.3	6			3	0	
-0.4						
-0.5	8		1			
-0.6		1				

Note. OL = Oral Language (k = 31 studies, n = 2,025 children); *EV* = Expressive Vocabulary (k = 20, n = 1,350); *RV* = Receptive Vocabulary (k = 23; n = 1,765); *AK* = Alphabet Knowledge (k = 13, n = 1,170); *PS* = Phonological Sensitivity (k = 13, n = 1,105); *OA* = Orthographic Awareness (k = 9; n = 880) implemented the interactive reading intervention in preschools, versus ten in kindergarten classrooms. No statistically significant age differences were detected on oral language and phonological sensitivity outcomes ($Q_{Oral\ Language}(1) = .61$, $p > .05$; $Q_{Expressive\ Vocabulary}(1) = 2.44$, $p > .05$; $Q_{Receptive\ Vocabulary}(1) = .001$, $p > .05$; $Q_{Phonological\ Sensitivity}(1) = .82$, $p > .05$). Interestingly, a moderate effect of the intervention was found for the alphabetic knowledge of children in kindergarten classrooms ($Q(1) = 8.47$, $p < .01$; $k = 8$, $d = .53$, *95%* CI = .34, .72), whereas children in preschool showed no growth at all ($k = 5$, $d = -.03$, *95%* CI = -.35, .29). In Figure 4.2, this contrast is displayed graphically. When the sample was restricted to at-risk groups, we again found a significant age effect for alphabetic knowledge ($k_{\text{kindergarten}} = 6$, $k_{\text{preschool}} = 5$; $Q(1) = 22.25$, $p < .001$) but not for the other outcome variables. In short, only kindergarten children showed a statistically significant growth in alphabetic knowledge.

Figure 4.2

Paired Comparison Chart of Preschool versus Kindergarten Children on all Outcome Measures, with Significant Age Differences on Alphabet Knowledge.



Note. OL = Oral Language; *EV* = Expressive Vocabulary; *RV* = Receptive Vocabulary; *AK* = Alphabet Knowledge *;* *PS* = Phonological Sensitivity

Intervention Characteristics as Moderators

To explain the variability in effect sizes, moderator analyses were conducted for five intervention characteristics: The adult who carried out the intervention, group size, type of intervention program, activities in the control group, and duration of the intervention.

First, we tested what was more effective: to be read to by a teacher or an experimenter. The random effects models were significant for the oral language composite ($Q(1) = 4.24$, $p < .05$) and expressive vocabulary ($Q(1) = 6.02$, $p < .05$), implying that experimenters such as the researcher or a graduate student were more effective than teachers (Oral Language: $k_{teacher} = 16$, $d = .35$, *95%* CI = .08, .62; $k_{experimenter} = 15$, $d = .79$, *95%* CI = .47, 1.10; Expressive Vocabulary: $k_{teacher} = 11$, $d = .28$, *95%* CI = .14, .69; $k_{experimenter} = 15$, $d = 1.10$, *95%* CI = .59, 1.56). It should be noted, however, that all but one study reported about teachers who interacted with children in small ($k = 5$, $n = 184$) or large groups ($k = 10$; $n = 1{,}295$), whereas experimenters read to children one-to-one ($k = 5$; $n = 145$) or in small ($k = 6$; $n = 299$) and large groups ($k = 4$; $n = 80$). When we examined both moderators simultaneously to find out which combination would be most effective ($Q(4) = 12.36$, $p < .05$), experimenters interacting with individual children seemed to have the strongest impact on children's oral language skills ($d = 1.38$, *95%* CI = .86, 1.89), differing significantly from the small to moderate effects that teachers revealed by reading to small groups ($Q(1) = 15.61$, p<001; $d = .15$, *95%* CI = -.24, .54) and large groups ($Q(1) = 8.69$, $p < .01$; $d = .48$, *95%* CI = .19, .78), as well as experimenters in large groups ($Q(1) = 5.36$, $p < .05$; $d = .34$, *95%* CI = -.28, .95). Insofar as sufficient studies were available with expressive vocabulary as a dependent measure, results were similar (see Table 4.2). No statistically significant differences were found for phonological sensitivity ($k_{teacher\text{-}large\ groups} = 6$, $k_{experimenter\text{-}small\ groups} = 5$; $Q(1) = 4.59$, $p > .05$).

Furthermore, the studies could be divided into three categories: 8 studies ($n = 260$) implemented Dialogic Reading (DR) as developed by Whitehurst et al. (1988), 11 studies ($n = 411$) tested the effects of similar techniques without referring to the specific Dialogic Reading-format and were coded as interactive reading (IR), and another 12 studies ($n = 1{,}354$) included extra classroom activities to support the interactive reading sessions (IR+). For the oral language composite, the random effects analysis revealed a statistically significant difference among the groups ($Q(2) = 11.38$, $p < .01$). DR was the least effective program ($d = .24$, *95%* CI = -.17, .64), differing significantly from the rather strong effect revealed by IR ($Q(1) = 7.54$, $p < .01$; $d = 1.01$, *95%* CI = .64, 1.39). IR also differed significantly from the IR+ programs which had a surprisingly low average effect size ($Q(1) = 7.95$, $p < .01$; $d = .38$, *95%* CI = .10, .66). As can be seen in Table 4.2, the patterns were similar for expressive and receptive vocabulary. In contrast to Karweit and Wasik's hypothesis (1996), the additional classroom activities offered by IR+ programs did not improve children's language skills more than single interactive reading sessions. A new perspective opened up, however, when the adult that carried out the intervention was taken into account as an additional moderator. First, IR+ programs were implemented by teachers in all but 1 study, whereas children in the IR condition were read to by experimenters in 10 out of 11 studies. Second, in the DR program, half of the studies were executed by teachers ($k_{teacher} = 4$, $n = 128$; $k_{experimenter} = 4$, $n = 132$), who tended to be less effective than experimenters ($Q(1) = 3.06$, $p = .08$): Experimenters seemed to be moderately effective in eliciting gains in oral language skills whereas teachers trained in dialogic reading techniques did not reveal effects ($d_{teacher} = -.08$, *95%* CI = -.59, .43; $d_{experimenter} = .58$, *95%* CI = .04, 1.11). When only experimenters were selected ($k = 14$, $n = 465$), the effect size differences between DR and IR were no longer significant ($Q(1) = .78$, $p > .05$). Finally, statistically significant differences were found in experiment fidelity scores across programs ($F(2, 28) = 10.79$, $p < .001$), with IR scoring significantly higher in fidelity than the studies that tested the effects of DR and IR+ ($M_{\mathrm{IR}} = 6.00$, $SD = 2.37$; $M_{\mathrm{DR}} = 3.38$, $SD = 2.45$; $M_{\mathrm{IR+}} = 2.08$, $SD = 1.31$), implying that the better controlled experiments were implemented by experimenters.

Overall, the country and/or language did not explain any variability in the effects ($k_{English/US} = 21$, $n = 1{,}125$; $Q_{Oral\ Language}(1) = 1.14$, $p > .05$; $Q_{Alphabet\ Knowledge}(1) = .95$, $p > .05$; $Q_{Phonological\ Sensitivity}(1) = .001$, $p > .05$). Unfortunately, all experiments with control groups that received an intervention were conducted in the United States and in English ($k = 11$, $n = 675$), whereas studies that included control group children who received the standard school program were conducted in both English ($k = 10$, $n = 450$) and other languages such as Dutch, Hebrew, Portuguese, or Spanish ($k = 10$; $n = 900$). Significant group differences in the activity by the control group were present for all oral language outcomes ($Q_{Oral\ Language}(1) = 9.82$, $p < .01$; $Q_{Expressive\ Vocabulary}(1) = 9.08$, $p < .01$; $Q_{Receptive\ Vocabulary}(1) = 8.42$, $p < .01$). As can be seen in Table 4.2, studies that included a control group that was only pre- and posttested revealed significantly lower effect sizes for the oral language outcomes than studies in which the control-group children were part of a non language-related intervention. This suggests that more elegantly designed studies with a higher fidelity score revealed higher effect sizes ($t(30) = 5.02$, $p < .001$).

We used a 16-week cutoff to be close to an intervention of at least half a 10-month school year as the contrasts could not have been tested when we split at 5 months. Interestingly, children's oral language and alphabetic knowledge did not seem to be influenced by the duration of the interactive reading intervention ($Q_{Oral\ Language}(1) = 1.53$, $p > .05$; $Q_{Expressive\ Vocabulary}(1) = 1.67$, $p > .05$; $Q_{Recpetive\ Vocabulary}(1) = .27$, $p > .05$; $Q_{Alphabet\ Knowledge}(1) = .06$, $p > .05$), whereas phonological sensitivity skills significantly improved as the duration of the intervention increased ($Q(1) = 4.85$, $p < .01$). That is, interventions that were implemented during a short period ($M_{weeks} = 11.33$, $SD = 5.16$; $M_{sessions} = 27.17$, $SD = 6.40$) had a smaller effect on phonological sensitivity than interventions that were spread over 4 months to a school year ($k_{short} = 6$, $d = .21$, *95%* CI = -.04, .46; $k_{long} = 7$, $d = .60$, *95%* CI = .36, .83).

Table 4.2
Meta-Analytic Results per Outcome Measure

	k	*n*	*d*	95% CI	*Q* [a]	*p*
Oral Language						
Total Set	**31**	**2,025**	**.54*** **	**.33, .74**	**318.15**	**< .001**
Design					1.69	.19
Experiment	19	671	.66***	.38, .93		
Quasi-Experiment	12	1,354	.38*	.07, .69		
School Type					.61	.44
Preschool	21	1,030	.60***	.34, .85		
Kindergarten	10	995	.42*	.08, .77		
Experimenter * Group size					12.36	.02
Experimenter-Individual	5	145	1.38***	.86, 1.89		
Experimenter-Small Group	6	299	.59*	.13, 1.05		
Experimenter-Large Group	4	80	.34	-.28, .95		
Teacher-Small Group	5	184	.15	-.24, .54		
Teacher-Large Group	10	1,295	.48***	.19, .78		
Type of Intervention Program					11.38	.003
DR	8	260	.24	-.17, .64		
IR	11	411	1.01***	.64, 1.39		
IR+	12	1,354	.38**	.10, .66		
Program Type * Experimenter					3.06	.08
DR-Teacher	4	128	-.08	-.59, .43		
DR-Experimenter	4	132	.58*	.04, 1.11		
Activity Control Group					9.82	.002
Intervention	11	675	.95***	.63, 1.27		
No intervention	20	1,350	.34**	.12, .55		
Duration Intervention					1.53	.22
Short (<16 weeks)	20	1,002	.64***	.38, .89		
Long	11	1,023	.37*	.03, .70		
Expressive Vocabulary						
Total Set	**20**	**1,350**	**.62*** **	**.29, .95**	**212.00**	**.00**
Design					.41	.52
School Type					2.44	.12
Experimenter * Group size					8.02	.046
Experimenter-Individual	5	145	1.40**	.56, 2.23		
Experimenter-Small Group	4	223	.73	-.16, .62		
Teacher-Large Group	7	1,011	.47	-.08, 1.02		
Type of Intervention Program					9.29	.01
DR	7	243	.20	-.30, .70		
IR	6	253	1.36***	.78, 1.94		
IR+	7	1,011	.47	-.01, .95		
Activity Control Group					9.08	.003
Intervention	7	537	1.26***	.74, 1.77		
No intervention	13	970	.29	-.07, .65		
Duration Intervention					1.67	.197

	k	*n*	*d*	*95%* CI	*Q* [a]	*p*
Receptive Vocabulary						
Total Set	**23**	**1,765**	**.45*****	**.22, .68**	**121.91**	**< .001**
Design					.08	.78
School Type					.001	.98
Experimenter * Group size					1.15	.29
Type of Intervention Program					6.32	.04
DR	7	243	.17	-.24, .58		
IR	5	173	.98***	.49, 1.48		
IR+	11	1,295	.42**	.15, .69		
Activity Control Group					8.42	.004
Intervention	5	444	.92***	.56, .128		
No intervention	18	1,267	.31**	.11, .51		
Duration Intervention					.27	.61
Alphabet Knowledge						
Total Set	**13**	**1,170**	**.39****	**.16, .62**	**40.28**	**< .001**
Design					.02	.88
School Type					8.47	.004
Preschool	5	269	-.03	-.35, .29		
Kindergarten	8	901	.53***	.34, .72		
Duration Intervention					.06	.81
Phonological Sensitivity						
Total Set	**13**	**1,105**	**.43*****	**.25, .62**	**77.92**	**84.60**
Design					.81	.37
School Type					.82	.37
Experimenter * Group size					4.59	.10
Duration Intervention					4.85	.028
Short (<16 weeks)	6	413	.21	-.04, .46		
Long	7	699	.60***	.36, .83		
Orthographic Awareness						
Total Set	**9**	**880**	**.41****	**.20, .62**	**19.95**	**.01**

Note. k = number of studies; n = total number of participants; *95%* CI = confidence interval; [a]*Q* for subset stands for homogeneity ($df = k - 1$); *Q* for moderator stands for effects of contrasts (df = number of subsets – 1). Contrasts were not tested when k<4 studies. Except for the Oral Language Composite, point estimates were not presented when the $Q_{between}$ was not significant ($p > .05$); * $p < .05$, ** $p < .01$, ***$p < .001$

Discussion

This meta-analysis tested the effects of an intervention that is thought to enhance the quality of adult-child storybook reading in early education and is expected to foster children's language and literacy development as a consequence. Results showed that children's oral language as well as print knowledge benefited from interaction before, during and after shared reading sessions. That is, about 6% of the growth in oral language skills could be explained by an interactive reading intervention in an educational setting (r = .25). Because the program focuses on techniques such as eliciting and reinforcing verbal responses by the child, it seems likely that children's expressive vocabulary skills will benefit most. Indeed, a moderate effect size was found for expressive vocabulary, explaining 8% of the variance (r = .28). These results indicate that the quality of book reading is important in addition to its frequency (Bus et al., 1995; Scarborough & Dobrich, 1994). The findings could not be attributed to differences in design characteristics or publication biases. When translated into a binominal effect size display or a change in success ratio, the oral language of children exposed to an interactive reading program gained 28% more than their peers in a control group, meaning that with interaction 64% improved in oral language, compared to 36% of children who were not part of the intervention. As in the medical domain when drugs are prescribed to millions of people because of a difference as small as 3% between the control and intervention groups (see Bus, 2001; Bus & Van IJzendoorn, 2004), it can be argued that interactive reading in early education warrants implementation.

Although adults were not trained to refer to print, 7% of the variance in kindergarten children's alphabetic knowledge could be explained by the interactive reading program. As expected, older children were able to significantly expand their emergent alphabetic knowledge, whereas younger children's print knowledge hardly benefited from interactive storybook encounters. Phonological sensitivity improved from interventions that were spread over a longer period of time, such as a school year. Following Lonigan (2006), we may define this result as a "dissociation effect" (p. 85). He demonstrated that an oral language intervention significantly affected measures of rhyme and blending. Albeit the studies do not provide data related to the qualities of the interactions during reading, it is conceivable that kindergarten teachers made more references to print than preschool teachers and/or that children with some knowledge of print may have elicited discussion of print features. Alternatively, a storybook itself might emphasize print and enhance print knowledge by varying font types and sizes, displaying some utterances in text balloons, or using rhyme and alliterations (Justice & Lankford, 2002). Unfortunately, hardly any information was provided about print-salient features within the storybooks that were used in the intervention studies. We speculate that children's ability to divide their attention between an adult and a book increases with growing experience in comprehending and interpreting a story's content. As children grow older, they might have control of skills to explore and process other features of the printed text, such as single letters, while listening to and interacting with an adult at the same time, whereas younger children need to invest all efforts in understanding the story. This assumption seems to contradict eye-tracking research by Evans and Saint-Aubin (2005) and Justice et al. (2005), but more recent studies by the same group of authors (Justice, Pullen, & Pence, 2008; Roy-Charland et al., 2007) show that the degree to which children learn about print during book reading depends on the extent that adults (non) verbally refer to print and the materials' characteristics. Assuming that the input of children and their environment affect each other reciprocally, we propose a transactional model of book reading to explain that not only does the interactive reading style of the teacher affect learning but the child's role is important as well (Sameroff & Fiese, 2000).

Differential effects of reading experience might also be reflected in the number of repetitions of the same book. Familiarity with the story content due to repeated readings of one story may create new opportunities to shift children's attention to other features of the text, as G. Phillips and McNaughton (1990) suggested based on a series of case studies. Unfortunately, these hypotheses could not be tested in the current meta-analysis, as almost all studies that included print-related measures reread a storybook at least once. Next to quality-related explanations, it seems likely that both language skills and print knowledge affect each other reciprocally (e.g., NICHD, 2005; Poe, Burchinal & Roberts, 2004; Samuelsson et al., 2007; Speece, Ritchey, Cooper, Roth, & Schatschneider, 2004; Storch & Whitehurst, 2002). The lexical restructuring hypothesis assumes that remembering and recognizing words in smaller segments become more efficient as children acquire more and more words via spoken language experiences (Metsala & Walley, 1998). On the other hand, the so-called Jabberwocky effect implies that phonological sensitivity stimulates vocabulary knowledge. As in Lewis Carroll's poem, phonemes carry meaning: Children who know more about units of words can use that knowledge to tune to parts of new words that have meaning for them and expand their vocabulary (Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003). Further research is needed, though, to understand the underlying processes that might explain the additional effect of interactive reading on print knowledge.

Based on a meta-analysis of 31 experiments that study different designs and populations, we can make some propositions for interventions that work best (Lipsey & Wilson, 2001). First, most explicit effect sizes appeared to be present in experiments that were highly controlled and executed by examiners. Teachers seemed to have difficulty with fostering the same growth in young children's language and literacy skills as researchers. It can be hypothesized that teachers were not successful in incorporating and internalizing the novel strategies in line with the intentions of the program developers (Dickinson & Sprague, 2001). Studying the factors that promote the transfer of evidence-based interventions to real-world settings is important, though. A cost-effective potential for generalization to future cohorts as well as the integration within the regular curriculum will widen the scope of the intervention (Aram, 2006). To promote treatment fidelity, it may be critical that the social component of the implementation process is emphasized with several opportunities for feedback and positive reinforcement next to the training in more technical aspects such as the theory behind the intervention (Rohrbach et al., 2006; Shernoff & Kratochwill, 2007). Even though the techniques developed by Whitehurst et al. (1988) are most tried and improved, the lowest effects were found for this program. It should be noted here that Dialogic Reading was implemented by showing and discussing a videotape in a session before and halfway through the program. Less standardized interactive reading interventions may incorporate more opportunities to coach teachers, discuss and solve concrete problems, and adapt the program to the needs of a specific classroom. Furthermore, in the current meta-analysis, teachers participated mostly in interventions that affected all kinds of classroom activities. It seems plausible that investing in play, art, or drama activities might have distracted teachers from giving as much attention to the interactive storybook reading as the researchers had anticipated, or as is evident for children who are part of a single interactive reading program. However, the above speculations provide only a partial solution to the interpretation of the effect sizes. The hypotheses can only be confirmed when experimenters and teachers are contrasted across program types. For instance, it should be tested whether lower effect sizes will also be found when experimenters implement IR+ programs or when teachers carry out an IR-only program. Varying the intensity of the intervention will be interesting as well, as will be exploring the frequency with which a book should be repeated to benefit optimally from the interaction. In sum, the current findings raise important questions to investigate through further research.

Second, we did not find support for Karweit and Wasik's (1996) conclusion that teachers should read to small groups when it is feasible to do so; neither did the results confirm Dickinson and Sprague's (2001) assumption that it is too challenging for teachers to tune to children's developmental level while reading to whole classrooms. In fact, this meta-analysis demonstrated that children's skills improved when their teachers engaged them in whole-group interactive reading sessions. Teachers might feel more inclined to focus their questions on events that are directly related to the book in order to keep control over the reading session and help the children to focus on the story's meaning and vocabulary as a consequence (Dickinson & Smith, 1994; Karweit & Wasik, 1996). In small group settings, the amount of extraneous talk and the opportunity for each child to elaborate on his or her own experiences might be distracting. Taken together, the quantitative summary of the research base thus far shows that interactive reading is worth implementing in classroom settings. To bridge the gap between research and practice, however, more research is needed to investigate the most promising implementation strategies.

Third, as groups at risk will be more susceptible to and in need of stimulation at school (Mol et al., 2008), they seem to be a promising group to invest in. All but four studies included children at risk for language and literacy impairments.

Cautions and Limitations

The result of a meta-analysis can be only as good as the studies that are included. Due to a lack of detail in the description of the intervention and its application in some studies, it was hard to extract data on all the characteristics that we were interested in. For instance, researchers reported the number of sessions they intended the teachers to offer to the children, but they did not seem to have observed whether this intensity was actually realized. We could not disentangle whether the duration of the intervention was a real confounder because the interactive reading programs with additional classroom activities were spread over a longer period than the dialogic reading interventions. Future studies should consider Lonigan and Whitehurst's (1998) approach: They presented separate and significantly different data for day care centers that reported to have held frequent and infrequent reading sessions, categorized as high- versus low-compliant centers. Overall, most studies seemed to lack control over the quality and frequency of book reading in control groups, especially when control-group children did not receive an intervention. This might restrict our conclusions regarding the unique effects of interactive reading in the classroom.

Because only half of the studies included a measure of print knowledge, most of the moderators could be analyzed only for oral language outcomes. Besides, studies that presented not raw posttest means and standard deviations but scores that took into account the pretest scores revealed stronger effects. Although none of the studies reported that their groups differed on the pretest, it seems tenable that a child's initial skills affect and account for his or her learning potential. Finally, the current findings are not completely independent from the meta-analysis on dialogic reading in parent-child dyads (see Mol et al., 2008). For two studies (Crain-Thoreson & Dale, 1999; Lonigan & Whitehurst, 1998), we included the same control-group posttest data as Mol et al. (2008).

Practical Implications

An interactive exposure to storybooks can be considered as an effective stimulant for the development of two pillars of learning to read: oral language and print knowledge. The current meta-analysis showed that interactive qualities of book reading in classrooms are effective supplements to book reading (Bus et al., 1995; Scarborough & Dobrich, 1994). Teachers who read to whole groups and accompanied the storybooks with extra activities knew to elicit moderate effects in oral language and print knowledge. The added value of interactive reading was reflected best in children who individually interacted with experimenters. Although the included studies did not provide enough details to grasp exactly what happened during the interactive reading sessions, it seems evident that children had a chance to learn about the story language as well as the written format of read-aloud texts. As program type and experimenter appeared to be interrelated in the current meta-analysis, more research is needed to disentangle the specific effects for interventions with and without additional activities that are implemented by experimenters versus teachers. Contrasting different implementation strategies to enhance effectiveness might be helpful as well. Compared to a short training by videotape, closely monitoring or coaching teachers might yield better opportunities to internalize a program's principles and adapt the trained techniques to the developmental level of the children in a classroom. Future observations are needed to explain our finding that interactive reading also affects children's print knowledge: To what extent do adults use the reading sessions to teach letters and sounds, do print-salient features attract attention, or do older children spontaneously pay attention to print and expand their skills by themselves when they grow more knowledgeable?

Appendix

Appendix 4.1

Study, Population, and Intervention Characteristics per Study Included in the Meta-Analysis.

First Author	*Year*	*Publ. Status* [a]	N_{tot} (n_{int} & n_{contr})[b]	*Design*[c]	*Country (Language)*	*School Type*[d]	*At Risk*	*Interv. Program*[e]	*Intervention ContrGr*	*Adult* ($n_{readers}$ *in Int.Gr)*	*Group Size* (n_{childr})	*Duration in weeks* ($n_{sessions}$)[f]	*Outcome*[g]	*ES* (d)[h]	*SE*
Aram-st1[i]	2006	Publ.	40	QE	Israel	PreS	Yes	IR+	No	Teacher	Small	28	*OL-RV*	0.47	0.23
			(20 & 20)		(Hebrew)					(6)	(4-6)	(M = 50)	*AK*	0.07	0.23
													PS	0.67	0.32
													OA	-0.21	0.23
Aram-st2[j]	2006	Publ.	38	QE	Israel	K	Yes	IR+	No	Teacher	Small	28	*OL-RV*	0.53	0.24
			(17 & 21)		(Hebrew)					(6)	(4-6)	(M = 50)	*AK*	0.61	0.24
													PS	1.48	0.37
													OA	0.59	0.24
Aram	2004	Publ.	59	QE	Israel	PreS	Yes	IR+	No	Experimenter	Small	28	*OL*	-0.36	0.27
			(35 & 24)		(Hebrew)					(2)	(4-6)	(M = 66)	*PS*	0.37	0.27
													OA	0.39	0.19
Brabham	2002	Publ.	78	E	U.S.	G1	No	IR	Yes	Teacher	Large	2	*OL*	1.02	0.12
			(39 & 39)		(English)					(5)	(6-9)	(R = 6)	*RV*	1.52	0.18
Crain-Thoreson	1999	Publ.	22	E	U.S.	PreS	Yes	DR	No	Teacher	Individual	8	*OL*	-0.13	0.25
			(13 & 9)		(English)					(8)		(R = 32)	*EV*	-0.18	0.31
													RV	-0.04	0.43
de Oliviera Fontes	2004	Publ.	38	E	Brasil	K	Yes	IR	No	Experimenter	Small	16	*OL*	0.84	0.17
			(19 & 19)		(Portuguese)					(1)	(2-4)	(M = 13.8)	*EV*	0.60	0.33
													RV	0.92	0.34
													AK	0.25	0.33
													PS	0.66	0.33
													OA	-0.10	0.32
Droop-st1[k]	2005	Publ.	95	QE	Netherlands	K	Yes	IR+	No	Teacher	Large	16	*OL*	0.66	0.15
			(42 & 53)		(Dutch)					(11)	(whole class)	(M = 32)	*EV*	0.71	0.21
													RV	0.62	0.21
													AK	0.66	0.21
													PS	0.29	0.09
													OA	0.85	0.22

Droop-st2[k]	2005	Publ.	168 (91 & 77)	QE	Netherlands (Dutch)	K	No	IR+	No	Teacher (11)	Large (whole class)	16 (M = 32)	*OL*	-0.13	0.11
													EV	-0.16	0.16
													RV	-0.10	0.15
													AK	0.08	0.15
													PS	0.02	0.07
													OA	0.35	0.16
Droop-st3[l]	2005	Publ.	248 (94 & 154)	QE	Netherlands (Dutch)	K	No	IR+	No	Teacher (17)	Large (whole class)	20 (M = 40)	*OL*	-0.13	0.09
													EV	-0.07	0.13
													RV	-0.19	0.13
													AK	0.64	0.13
													PS	0.28	0.06
													OA	0.71	0.13
Droop-st4[m]	2005	Publ.	81 (39 & 42)	QE	Netherlands (Dutch)	K	Yes	IR+	No	Teacher (17)	Large (whole class)	20 (M = 40)	*OL*	0.20	0.16
													EV	0.24	0.22
													RV	0.16	0.22
													AK	0.73	0.23
													PS	0.47	0.10
													OA	0.29	0.22
Droop-st5[k]	2005	Publ.	113 (62 & 51)	QE	Netherlands (Dutch)	K	Yes	IR+	No	Teacher (17)	Large (whole class)	20 (M = 40)	*OL*	0.35	0.14
													EV	0.15	0.19
													RV	0.54	0.19
													AK	0.60	0.19
													PS	0.53	0.09
													OA	0.52	0.19
Karweit-st1[i]	1989	Publ.	86 (43 & 43)	QE	U.S. (English)	PreS	Yes	IR+	No	Teacher (2)	Large (whole class)	28 (R = 56)	*OL*	0.47	0.11
													RV	0.52	0.22
Karweit-st2[j]	1989	Publ.	120 (60 & 60)	QE	U.S. (English)	K	Yes	IR+	No	Teacher (2)	Large (whole class)	28 (R = 56)	*OL*	0.46	0.09
													RV	0.24	0.18
													AK	0.62	0.19
													PS	1.04	0.19
Kertoy	1994	Publ.	16 (8 & 8)	E	U.S. (English)	K	No	IR	Yes	Experimenter (2)	Individual	3 (R = 4)	*OL-EV*	0.54	0.51

First Author	*Year*	*Publ. Status* [a]	*N_{tot} (n_{int}&n_{contr})* [b]	*Design* [c]	*Country (Language)*	*School Type* [d]	*At Risk*	*Interv. Program* [e]	*Interven tion ContrGr*	*Adult ($n_{readers}$ in Int.Gr)*	*Group Size (n_{childr})*	*Duration in weeks ($n_{sessions}$)* [f]	*Outcome* [g]	*ES (d)* [h]	*SE*
van Kleeck	2006	Publ.	30 (15 & 15)	E	U.S. (English)	PreS	Yes	IR	No	Experimenter (*N* not rep.)	Individual	8 (*R* = 16)	*OL*	1.18	0.23
													EV	1.07	0.28
													RV	1.41	0.41
Lamb-st1[n]	1986	Diss.	14 (5 & 9)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (1)	Large (10)	10 (*M* = 50)	*OL*	-0.21	0.40
													RV	-0.01	0.56
													AK	-0.26	0.56
Lamb-st2[o]	1986	Diss.	13 (5 & 8)	E	U.S. (English)	PreS	Yes	IR	No	Experimenter (1)	Large (10)	10 (*M* = 50)	*OL*	-0.08	0.41
													RV	0.16	0.57
													AK	0.92	0.60
Lonigan-st1[p]	1998	Publ.	27 (16 & 11)	E	U.S. (English)	PreS	yes	DR	No	Teacher (*N* not rep.)	Small (5)	6 (*M* = 6.8)	*OL*	-0.58	0.23
													EV	-0.61	0.28
													RV	-0.51	0.40
Lonigan-st2[q]	1998	Publ.	31 (15 & 16)	E	U.S. (English)	PreS	Yes	DR	No	Teacher (*N* not rep.)	Small (5)	6 (*M* = 17.4)	*OL*	0.25	0.21
													EV	0.34	0.26
													RV	0.06	0.36
Lonigan-st1[r]	1999	Publ.	46 (17 & 29)	E	U.S. (English)	PreS	Yes	DR	Yes	Experimenter (*N* not rep.)	Small (2-5)	6 (*M* = 18.5)	*OL*	0.19	0.18
													EV	0.06	0.22
													RV	0.46	0.31
													PS	-0.30	0.16
Lonigan-st2[s]	1999	Publ.	49 (17 & 32)	E	U.S. (English)	PreS	Yes	DR	No	Experimenter (*N* not rep.)	Small (2-5)	6 (*M* = 18.5)	*OL*	0.04	0.17
													EV	0.04	0.21
													RV	0.03	0.30
													PS	0.14	0.16
Mautte-st1[n]	1990	Diss.	29 (9 & 20)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (1)	Large (7)	20 (*M* = 60)	*OL*	0.34	0.40
Mautte-st2[o]	1990	Diss.	24 (9 & 15)	E	U.S. (English)	PreS	Yes	IR	No	Experimenter (1)	Large (7)	20 (*M* = 60)	*OL*	1.46	0.40
Morrow-st1[t]	1988	Publ.	39 (25 & 14)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (*N* not rep.)	Individual	10 (*M* = 10)	*OL-EV*	2.71	0.47
Morrow-st2[u]	1988	Publ.	40 (27 & 13)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (*N* not rep.)	Individual	10 (*M* = 10)	*OL-EV*	0.93	0.43

Morrow	1989	Publ.	90 (45 & 45)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (*N* not rep.)	Small (3)	11 (*M* = 11)	*OL-EV*	2.27	0.45
Murphy	2007	Diss.	17 (8 & 9)	E	U.S. (English)	PreS	Yes	DR	Yes	Experimenter (1)	Small (3)	8 (*M* = 23)	*OL*	0.69	0.35
													AK	0.60	0.50
													PS	1.48	0.27
Valdez-Menchaca	1992	Publ.	20 (10 & 10)	E	Mexico (Spanish)	PreS	Yes	DR	No	Experimenter (1)	Individual	7 (*M* = 30)	*OL*	1.57	0.40
													EV	1.75	0.26
													RV	1.21	0.50
Wasik	2001	Publ.	121 (61 & 60)	QE	U.S. (English)	PreS	Yes	IR+	Yes	Teacher (2)	Large (whole class)	15 (*R* = 45)	*OL*	1.35	0.49
													EV	2.06	0.12
													RV	1.08	0.30
Wasik	2006	Publ.	185 (124 & 61)	QE	U.S. (English)	PreS	Yes	IR+	Yes	Teacher (10)	Large (whole class)	36 (*R* = 72)	*OL*	0.61	0.14
													EV	0.45	0.16
													RV	0.78	0.11
													AK	-0.33	0.22
Whitehurst	1994	Publ.	48 (26 & 22)	E	U.S. (English)	PreS	Yes	DR	No	Teacher (*N* not rep.)	Small (5)	6 (*M* = 16.6)	*OL*	0.10	0.16
													EV	0.09	0.15
													RV	0.15	0.15

Notes. [a] Publ. Status: Publ. = Published, Diss. = Dissertation; [b] Sample sizes were adapted if studies included more than one intervention or control group and all groups met the inclusion criteria. [c] Design: QE = Quasi-Experiment, E = Experimental Study; [d] School Type at the start of the study: PreS = Preschool, K = Kindergarten, G1 = Grade 1; [e] Type of Intervention Program: DR = Dialogic Reading, IR = Interactive Reading, IR+ = Interactive Reading with Additional Activities; [f] Duration intervention in weeks: the number of sessions is presented between brackets. M = Mean number of sessions as reported in the article; R = Recommended number of sessions as proposed in the article when no observed Mean is provided; In the analyses, interventions less than 16 weeks, were treated as "short"; [g] Outcome: OL = Oral Language Composite, RV = Receptive Vocabulary, OL-EV = Oral Language Measure represents Expressive Vocabulary, OL-RV = Oral Language Composite reflects Receptive Vocabulary, AK = Alphabetic Knowledge, PS = Phonological Sensitivity, OA = Orthographic Awareness; [h] Effect Sizes were based on posttest mean and standard deviations of an experimental and control group, weighted by sample size. A positive sign indicates a favorable outcome for the experimental group; [i] Preschool (sub)sample; [j] Kindergarten (sub)sample; [k] Second Language Learners, at risk; [l] First Language Learners, not at risk; [m] First Language Learners, at risk; [n] interactive reading experimental group versus a typical reading control group; [o] interactive reading experimental group versus a no-intervention control group; [p] Low-compliance centers; [q] High-compliance centers; [r] Dialogic Reading experimental group versus a typical reading control group; [s] Dialogic Reading experimental group versus a no-intervention control group; [t] repeated reading experimental group versus control group; [u] different books reading group versus control group.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

Adams, M. J. (1990). *Beginning to read.* Cambridge, MA: MIT Press.

* Aram, D. (2006). Early literacy interventions: The relative roles of storybook reading, alphabetic activities, and their combination. *Reading and Writing, 19,* 489-515.

* Aram, D., & Biron, S. (2004). Joint storybook reading and joint writing interventions among low SES preschoolers: Differential contributions to early literacy. *Early Childhood Research Quarterly, 19,* 588-610.

Ard, L. M., & Beverly, B. L. (2004). Preschool word learning during joint book reading: Effect of adult questions and comments. *Communication Disorders Quarterly, 26,* 17-28.

Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129,* 195-215.

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *Elementary School Journal, 107,* 251-273.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98,* 44-62.

Blok, H. (1999). Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning, 49,* 343-371.

Bodrova, E., & Leong, D. J. (2006). Vygotskian perspectives on teaching and learning early literacy. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 243-256). New York: The Guilford Press.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2* [Computer Software]. Englewood, NJ: Biostat.

* Brabham, E. G., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94,* 465-473.

Bus, A. G., & van IJzendoorn, M. H. (1988). Mother-child interactions, attachment, and emergent literacy: A cross-sectional study. *Child Development, 59,* 1262-1272.

Bus, A. G., & van IJzendoorn, M. H. (1995). Mothers reading to their three-year-olds: The role of mother-child attachment security in becoming literate. *Reading Research Quarterly, 30*, 998-1015.

Bus, A. G., & van IJzendoorn, M. H. (2004). Meta-analysis in reading research. In N. Duke & M. Mallette (Eds.), *Literacy research methods* (pp. 227–252). New York: Guilford Press.

Bus, A. G. (2001). Parent-child book reading through the lens of attachment theory. In L. Verhoeven & C. E. Snow (Eds.), *Literacy and Motivation; Reading Engagement in Individuals and Groups* (pp. 39-53). Mahwah: Lawrence Erlbaum Associates.

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1-21.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Collins, M. F. (2005). IRA Outstanding Dissertation Award for 2005: ESL preschoolers' English vocabulary acquisition from storybook reading. *Reading Research Quarterly, 40,* 406-408.

* Crain-Thoreson, C., & Dale, P. S. (1999). Enhancing linguistic performance: Parents and teachers as book reading partners for children with language delays. *Topics in Early Childhood Special Education, 19,* 28-39.

* de Oliviera Fontes, M. J., & Cardoso-Martins, C. (2004). Efeitos da leitura de histórias no desenvolvimento da linguagem de crianças de nível sócio-econômico baixo [Effects of story reading on low SES children's language development]. *Psicologia: Reflexao e Crítica, 17,* 83-94.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318,* 1387-1388.

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E.-S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology, 95,* 465-481.

Dickinson, D. K., & Smith, M. W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly, 29,* 104-122.

Dickinson, D. K., & Sprague, K. E. (2001). The nature and impact of early childhood care environments on the language and early literacy development of children from low-income families. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of Early Literacy Research* (Vol. 1, pp. 263-280). New York: The Guilford Press.

* Droop, M., Peters, S., Aarnoutse, C., & Verhoeven, L. (2005). Effecten van stimulering van beginnende geletterdheid in groep 2 [Effects of beginning literacy education in kindergarten]. *Pedagogische Studiën, 82,* 160-180.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method for accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89-98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455-463.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24,* 174-187.

Evans, M. A., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading - Evidence from eye movement monitoring. *Psychological Science, 16,* 913-920.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., Braun, M., & O'Connor, R. E. (2001). Is reading important in reading-readiness programs? A randomized field trail with teachers as program implementers. *Journal of Educational Psychology, 93,* 251-267.

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15,* 75-90.

Hart, B., & Risley, T. (2003). The early catastrophe: The 30 million word gap. *American Educator, 27,* 4-9.

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11,* 49-76.

Hedges, L. V. (2007). Correcting a significance test for clustering. *Journal of Educational and Behavioral Statistics, 32,* 151-179.

Ioannidis, J., Patsopoulos, N., & Evangelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. *BMJ, 335,* 914-916.

Juel, C. (2006). The impact of early school experiences on initial reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 410-426). New York: The Guilford Press.

Justice, L. M., & Ezell, H. K. (2002). Use of storybook reading to increase print awareness in at-risk children. *American Journal of Speech-Language Pathology, 11,* 17-29.

Justice, L. M., & Lankford, C. (2002). Preschool children's visual attention to print during storybook reading: Pilot findings. *Communication Disorders Quarterly, 24,* 9-19.

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools, 36,* 17-32.

Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Developmental Psychology, 44,* 855-866.

Justice, L. M., Skibbe, L., Canning, A., & Lankford, C. (2005). Pre-schoolers, print and storybooks: An observational study using eye movement analysis. *Journal of Research in Reading, 28,* 229-243.

* Karweit, N. (1989). The effects of a story-reading program on the vocabulary and story comprehension skills of disadvantaged prekindergarten and kindergarten students. *Early Education and Development, 1,* 105-114.

Karweit, N., & Wasik, B. A. (1996). The effects of story reading programs on literacy and language development of disadvantaged preschoolers. *Journal of Education for Students Placed at Risk, 1,* 319-348.

* Kertoy, M. K. (1994). Adult interactive strategies and the spontaneous comments of preschoolers during joint storybook readings. *Journal of Research in Childhood Education, 9,* 58-67.

* Lamb, H. A. (1986). *The effects of a read-aloud program with language interaction.* Unpublished doctoral dissertation, The Florida State University, Tallahassee.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lonigan, C. J. (2006). Conceptualizing phonological processing skills in prereaders. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 77-89). New York: The Guilford Press.

* Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention, 22,* 306-322.

* Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13,* 263-290.

Mason, J. M. (1992). Reading stories in preliterate children: A proposed connection to reading. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 215-241). Hillsdale, NJ: Lawrence Erlbaum.

* Mautte, L. A. (1990). *The effects of adult-interactive behaviors within the context of repeated storybook readings upon the language development and selected prereading skills of prekindergaten at-risk students.* Unpublished doctoral dissertation, University of South Florida, Tampa.

McCormick, C., & Mason, J. M. (1986). *Use of little books at home: A minimal intervention strategy that fosters early reading* (Techn Rep. No. 388). Urbana: University of Illinois, Center for the Study of Reading.

McKeown, M. G. & Beck, I. L. (2006). Encouraging young children's language interactions with stories. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 281-294). New York: The Guilford Press.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In L. C. Ehri & J. L. Metsala (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ: Lawrence Erlbaum.

Mol, S. E., Bus, A. G., De Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education and Development, 19,* 7-26.

Morris, D., Bloodgood, J. W., Lomax, R. G., & Perney, J. (2003). Developmental steps in learning to read: A longitudinal study in kindergarten and first grade. *Reading Research Quarterly, 38,* 302-328.

* Morrow, L. M. (1988). Young children's responses to one-to-one story readings in school settings. *Reading Research Quarterly, 23,* 89-107.

* Morrow, L. M. (1989). The effect of small group story reading on children's questions and comments. In: S. McCormick & J. Zuttell (Eds.), *Cognition and social perspectives for literacy research and instruction* (pp. 77-86). Chicago: 38th Yearbook of the National Reading Conference.

Morrow, L. M., & Smith, J. K. (1990). The effects of group size on interactive storybook reading. *Reading Research Quarterly,* 213-231.

Moschovaki, E., & Meadows, S. (2004). A short term longitudinal study of classroom book reading in Greek kindergarten schools. *Educational Studies in Language and Literature, 4,* 151-168.

Murphy, M. M. (2007). Enhancing print knowledge, phonological awareness, and oral language skills with at-risk preschool children in Head Start classrooms. Unpublished doctoral dissertation, University of Nebraska, Lincoln.

NICHD Early Child Care Research Network (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41,* 428 - 442.

Nielsen, D. C. (1993). The effects of four models of group interaction with storybooks on the literacy growth of low-achieving kindergarten children. In: D. J. Leu & C. K. Kinzer (Eds.), *Examining central issues in literacy research, theory, and practice* (pp. 279-287). Chicago: 42nd Yearbook of the National Reading Conference.

Ninio, A. (1980). Picture-book reading in mother-infant dyads belonging to two subgroups in Israel. *Child Development, 51,* 587-590.

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94,* 23-33.

Phillips, G., & McNaughton, S. (1990). The practice of storybook reading to preschool children in mainstream New Zealand families. *Reading Research Quarterly*, 25, 196-212.

Phillips, L. M., Norris, S. P., Mason, J. M., & Kerr, B. M. (1990). Effect of early literacy intervention on kindergarten achievement. In: J. Zuttell & S. McCormick (Eds.), *Literacy theory and research: Analyses from multiple paradigms* (pp. 199-207). Chicago: 39th Yearbook of the National Reading Conference..

Poe, M. D., Burchinal, M. R., & Roberts, J. E. (2004). Early language and the development of children's reading skills. *Journal of School Psychology, 42,* 315-332.

Raikes, H., Pan, B. A., Luze, G., Tamis-LeMonde, C. S., Brooks-Gunn, J., Banks Tarullo, J. C., et al. (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development, 6,* 193–235.

Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology, 35,* 20-28.

Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type II translation; Transporting prevention interventions from research to real-word settings. *Evaluation & the Health Professions, 29,* 302-333.

Roy-Charland, A., Saint-Aubin, J., & Evans, M. A. (2007). Eye movements in shared book reading with children from kindergarten to Grade 4. *Reading and Writing, 20,* 909-931.

Sameroff, A. J., & Fiese, B. H. (2000). Models of development and developmental risk. In C. H. Zeanah (Ed.), *Handbook of infant mental health* (pp. 3-19). New York: Guilford.

Samuelsson, S., Olson, R., Wadsworth, S., Corley, R., DeFries, J. C., Willcutt, E. et al. (2007). Genetic and environmental influences on prereading skills and early reading and spelling development in the United States, Australia, and Scandinavia. *Reading and Writing, 20,* 51-75.

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review, 14,* 245-302.

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language, 24,* 123-138.

Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73,* 445-460.

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87,* 218-229.

Shernoff, E. S., & Kratochwill, T. R. (2007). Transporting an evidence-based classroom management program for preschoolers with disruptive behavior problems to a school: An analysis of implementation, outcomes, and contextual variables. *School Psychology Quarterly, 22,* 449-472.

Speece, D. L., Ritchey, K. D., Cooper, D. H., Roth, F. P., & Schatschneider, C. (2004). Growth in early reading skills from kindergarten to third grade. *Contemporary Educational Psychology, 29,* 312-332.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934-947.

* Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading: A systematic extension to Mexican day care. *Developmental Psychology, 28,* 1106-1114.

van Elsäcker, W., & Verhoeven, L. (1997). Kleuters leren meer van voorlezen in kleine groepen [Children in kindergarten learn more from interactive reading in small groups]. *Pedagogische Studiën, 74,* 117-129.

* van Kleeck, A., Vander Woude, J., & Hammett, L. (2006). Fostering literal and inferential language skills in Head Start preschoolers with language impairment using scripted book-sharing discussions. *American Journal of Speech-Language Pathology, 15,* 85-95.

Viechtbauer, W. (2007). Confidence intervals for the amount of heterogeneity in meta-analysis. *Statistics in Medicine, 26,* 37-52.

* Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology, 93,* 243-250.

* Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology, 98,* 63-74.

* Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30,* 679-689.

Whitehurst, G. J., Epstein, J. N., Angel, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology, 86,* 542-555.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., & Fischel, J. E. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24,* 552-559.

Whitehurst, G. J., Zevenbergen, A. A., Crone, D. A., Schultz, M. D., Velting, O. N., & Fischel, J. E. (1999). Outcomes of an emergent literacy intervention from Head Start through second grade. *Journal of Educational Psychology, 91,* 261-272.

Yaden, D. B., Smolkin, L. B., & Conlon, A. (1989). Preschoolers' questions about pictures, print convention and story text during reading aloud at home. *Reading Research Quarterly, 24,* 188-214.

5

General Discussion

Parts of this chapter are based on:
Bus, A. G., Van IJzendoorn, M. H., & Mol, S. E. (in press). Meta-Analysis. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methods, volume 2*. New York: Guilford Press.

In his comprehensive book "*Progress in understanding reading*", Keith Stanovich (2000) states that we make progress by accumulating evidence from a host of interlocking studies, each of which may be of fairly low diagnosticity but that, taken together, present a coherent picture and warrant firm conclusions. He postulates: "We are a science that is custom-made for meta-analysis" (p. 3). The three comprehensive meta-analyses in this thesis are among approximately 120 meta-analyses that are conducted within the field of reading research thus far.

Hereafter, I will embed the findings of our meta-analyses in the broad variety of meta-analyses that have been published until September 2010. I will also present a meta-analysis of meta-analyses on interventions that aim to enhance the development of reading abilities in children. What do we know thus far and, specifically, how can we increase our understanding of the role of book reading in reading development from infancy to early adulthood?

A Brief History of Meta-Analysis

A century ago, Karl Pearson (1904) reported on one of the first meta-analytic combinations of the outcomes of a set of medical studies, and during the past few decades the approach became extremely popular in the so-called evidence-based medical science. It was the educational researcher Glass (1976) who coined the concept "meta-analysis" some 25 years ago and introduced it into the educational and behavioral science. He provided one of the most controversial examples of its application on psychotherapy studies, arguing that, in general, psychotherapy had considerable effect but that no specific treatment modality stood out (Smith & Glass, 1977).

To our knowledge, one of the first meta-analyses in reading was conducted by Kavale on correlates of reading: visual perceptual skills, auditory perceptual skills, and auditory-visual integration. He simply provided average correlations across studies between these predictors and success or failure in reading (Kavale, 1980, 1981, 1982; Kavale & Forness, 2000). During the past 15 years, meta-analysis has become widely used and hotly disputed in educational science. In fact, it seems that it has been applied on a much wider scale in education than in any other social or behavioral science. The reason may be that educational policy decisions (such as medical decisions) are supposed to be based on a firm foundation of empirical data (Slavin, 2002). Every decade the number of scientific papers is doubling (Garfield, 1979), and it becomes impossible even for the specialists – let alone the policymakers and practitioners – to keep track of the literature in their own field. More importantly, meta-analyses are increasingly being used to monitor new developments in any area of the social and behavioral sciences (Sutton & Higgins, 2008).

In the past, narrative reviews were considered the royal road to the synthesis of literature, and some narrative reviews indeed were very powerful in shaping the future of a field of inquiry (e.g., Adams, 1990). In a narrative review of high standards, the author tries to make sense of the literature in a systematic and, at the same time, creative way. In formulating a hypothesis for review in a precise manner, and in collecting systematically the pertinent papers to address the issue, the narrative reviewer does not act much differently from the meta-analyst. It is in the stage of data analysis that the narrative and meta-analytic reviewer go separate ways. Narrative reviewers may have the focus of telling readers what the field has and has not investigated more than what has been found. Insofar as they focus on conceptual analysis of studies, these might not include numerical results at all – as in a review of ethnographies of home literacy practices in different communities. The meta-analysts, on the contrary, proceed in a statistically rigorous way, analyzing studies that include numerical results. Effect sizes, quantitative indexes of relations among variables, are used to compare and communicate the strength of the summarized research findings (Hedges, 2008).

Cooper and Rosenthal (1980) showed experimentally that narrative reviewers are more inclined to commit type II errors (i.e., they tend to not reject the null hypothesis although it should be rejected on statistical grounds). Cooper and Rosenthal asked 41 graduate students and senior researchers to review a set of seven studies on the association between sex and persistence in performing rather dull tasks. Half of the reviewers were randomly assigned to a course on meta-analysis. Seventy-three percent of the untrained narrative reviewers found no association; only 32% of the meta-analysts came to this conclusion. The correct outcome was that female participants are significantly more persistent in performing boring tasks than males. In particular, in cases in which studies show insignificant trends, the accumulated effect size across these studies tends to be underestimated. Besides, narrative reviews are also more vulnerable to psychological factors. Bushman and Wells (2001) had 280 undergraduate students review 20 fictional studies, of which the salience of the title and serial order were manipulated. Interestingly, salient titles for the positive results led to overestimates of the actual relation, whereas salient titles for the negative results led to an underestimation of the effect magnitude (Bushman & Wells, 2001). It should be noted that despite this potential bias, narrative reviews remain indispensable, in particular in those areas in which a restricted number of empirical studies have been conducted or in the absence of strong research programs that unify the empirical approaches and make them comparable for meta-analytic purposes. Researchers sometimes persist in conducting a meta-analysis even when the exhaustive literature search results in the inclusion of only two or three studies (e.g., Jeynes, in press; Sénéchal & Young, 2008; Torgerson, Porthouse, & Brooks, 2003; Zucker, Moody, & McKenna, 2009).

On the other hand, not all research domains are ready for meta-analysis despite numerous studies. For example, the National Reading Panel (NRP) did not succeed in finding sufficient studies to meta-analyze effects of all formal efforts to increase the amounts of independent or recreational reading that children engage in, including sustained silent reading programs, because of a lack of studies that meet NRP standards such as experimental or quasi-experimental designs, including a control group (NRP, 2000). They concluded that it would be difficult to interpret the small collection of studies that remained as representing clear evidence that encouraging students to read more actually improves reading achievement. Only three of the 14 remaining studies reported any clear reading gains from encouraging students to read. However, one may wonder to what extent the selection criteria were responsible for this (counterintuitive) result. The selection of studies did not include a screening of studies in order to ensure that the participants needed what the treatment was designed to influence. The NRP routinely selected and analyzed studies that experimentally tested the efficacy of encouraging students to read more without ensuring that the participants in the selected studies indeed did not have the ability and opportunity outside of school to read independently (cf. Cunningham, 2001). Interestingly, the number of studies that correlated leisure-time reading activities to students' reading abilities is largely sufficient to synthesize quantitatively. In chapter 2 of this thesis, we included 40 studies targeting children attending grades 1 to 12 and 30 studies targeting undergraduate and graduate students. The correlation between leisure-time reading and students' reading comprehension and technical reading and spelling skills became stronger with age, which seems to be in line with a model of reciprocal causation. More skilled readers are more likely to choose to read more frequently which, in turn, will improve their reading abilities, whereas poor readers may not succeed in comprehending text, become less eager to read, and as a result, show stagnation in their reading development (Stanovich, 1986).

State of the Art in Meta-Analyses on Reading

We applied a computer search of PsycInfo and ISI, with the key words *literacy and meta-analysis* and *reading and meta-analysis* to trace relevant meta-analyses in the field of reading research. After excluding book chapters and dissertations, a relevant set of about 120 meta-analyses on reading resulted (see Appendix 5.1 for a summary of the about 100 reviews we could trace). Assuming that since 1966 approximately 130,000 research studies on reading have been conducted, with perhaps another 15,000 appearing before that time (NRP, 2000), only a small part of all available studies is meta-analyzed. The 120 meta-analyses in the reading domain upto 2010 cover at most 10% of all available studies on reading.

Most meta-analyses on reading synthesize the results of intervention studies. In an attempt to settle an ongoing debate on the best method to teach beginning reading skills, studies contrast whole language with basals (Stahl & Miller, 1989; Jeynes & Littell, 2000), systematic phonics instruction with no or incidental instruction in phonics (Ehri, Nunes, Stahl, & Willows, 2001), or reading instruction in the first or second language for bilingual children (Greene, 1997; Rolstad, Mahoney, & Glass, 2005; Slavin & Cheung, 2005). Other studies synthesize effects of special measures: programs to instruct phonemic awareness (Bus & van IJzendoorn, 1999; Ehri, Nunes, Willows, et al., 2001), guided oral reading (NRP, 2000), book reading in groups (Blok, 1999; chapter 4 of this thesis), question generation (NRP, 2000), repeated reading (Therrien, 2004), reading engagement (Guthrie, McRae, & Klauda, 2007), classroom discussion (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009), or learning to derive word meaning from context (Fukkink & De Glopper, 1998). Furthermore, it is evaluated how direct and strategy instructions support groups with learning disabilities (Edmonds et al., 2009; Sencibaugh, 2007; Swanson & Sachse-Lee, 2000), and whether one-to-one tutoring in reading (D'Agostino & Murphy, 2004; Elbaum, Vaughn, Hughes, & Moody, 2000; Ritter, Barnett, Denny, & Albin, 2009; Torgerson, King, & Sowden, 2002) or instruction in small groups especially stimulates these children's reading development (Elbaum, Vaughn, Hughes, & Moody, 1999; Elbaum, Vaughn, Hughes, Moody, & Schumm, 2000). Few studies test effects of school organization on reading achievement: class size (McGiverin, Gilman, & Tillitski, 1989) or summer holiday (Cooper, Nye, Charlton, & Lindsay, 1996). The effectiveness of children's learning experiences outside the classroom are examined by evaluating studies on parent involvement (Sénéchal & Young, 2008; NELP, 2008) and out-of-school programs (Cooper, Charlton, Valentine, & Muhlenbruck, 2000; Lauer et al., 2006). Only a few studies focus on interventions in the preschool ages and test effects of book reading in the family (Bus, van IJzendoorn, & Pellegrini, 1995; chapter 3, this thesis; NELP, 2008) or preschool intervention programs (Goldring & Pressbrey, 1986; Leseman, Otter, Blok, & Deckers, 1998, 1999; NELP, 2008; Manz, Hughes, Barnabas, Bracaliello, & Ginsburg-Block, in press; Marulis & Neuman, 2010; Piasta & Wagner, 2009). Recently, the increasing number of single studies that explore the opportunities of the computer for language instruction made some (preliminary) meta-analyses possible (Blok, Oostdam, Otter, & Overmaat, 2002; Moran, Ferdig, Pearson, Wardrop, & Blomeyer, 2008; Torgerson & Elbourne, 2002; Zucker et al., 2009).

Meta-Analyses about Meta-Analyses

Meta-analyses have not been positioned in a more crucial role than any other systematic form of inquiry. Meta-analyses are part of a series of connected steps in the description and explanation of human behavior that never reaches a final point (van IJzendoorn, 1994). Because meta-analyses are based on numerous decisions about collecting, coding, and analyzing the pertinent studies, meta-analytic results, in their turn, need to be replicated as well (Lytton, 1994). Even if replications of meta-analyses yield the same results, the most fruitful meta-analyses will lead to new hypotheses for further primary study (Eagly & Wood, 1994). By combining the results of several meta-analyses, researchers are able to construct models of associations between theoretically important variables which are not yet combined in any separate empirical study, and to show at what point the model still is incomplete.

A relatively small set of meta-analyses that we traced report about effects of instruction on reading comprehension (n = 19) and on word recognition (n = 11). From the stem-and-leaf display (see Figure 5.1), it appears that both word recognition and reading comprehension are susceptible to specific forms of instruction. Insofar as several dependent measures were available, we selected tests with established (by the experimenter or someone else) construct validity and reliability (using multiple measures of reliability) above experimenter tests. When a series of word-recognition outcomes were reported, we left out outcomes for selected words (e.g., pseudo- or only regularly spelled words).

For both word recognition and reading comprehension, outcomes are homogeneous according to an analysis on this data set with Comprehensive Meta-Analysis (Statistical Solutions Limited), even though the interventions cover a variety of instructions varying in form (group vs. one-to-one tutoring) and ranging from phonemic awareness to deriving meaning from context. Only a meta-analysis on the effects of reading comprehension interventions revealed outlying results (d = 1.23; Edmonds et al., 2009). As outlined in the introduction (see step 1), a rather diverse mix of intervention types might result in summary effects that are hard to interpret theoretically.

Another notable result is that effect sizes for word recognition skills exceed those for reading comprehension. With word recognition as a dependent variable, the median effect size of interventions is about half a standard deviation. With reading comprehension as the dependent measure, it is about a third of a standard deviation. These outcomes are similar whatever the focus of the study: improving word recognition, practicing comprehension skills, or one-to-one tutoring. In other words, word recognition is more susceptible to instruction than text comprehension. Reading comprehension is more strategic and based on higher level skills and may, therefore, be less trainable than decoding that is based on low-level skills. Interventions that include strategic and other higher level processes promise progress in comprehension (Pressley & Harris, 1994), but not to the same extent as a training of lower level skills warrants progress in word recognition. Because the interventions varied so much, we were unable to test characteristics of instruction. For instance, assuming that instruction on comprehension supports skills beyond those stimulated by word recognition, one may expect that the effect of comprehension instruction on comprehension is quite a bit higher than the effect of word-recognition instruction on comprehension, particularly after the early grades.

Figure 5.1
Stem-and-Leaf Display of d*-Indexes for Effects of Interventions on Achievement Test Scores in Word Recognition and Reading Comprehension.*

Word Recognition		Reading Comprehension
	1.2	3
	1.1	8
	1.0	
	.9	15
2	*.8*	
5	*.7*	2
	.6	7
7554	*.5*	
10	*.4*	133
2	*.3*	01245666
70	*.2*	58
	.1	
	0	

Note. Combine the stem (.1, .2, .3, etc.) with the leaves to the left and to the right to find *d* values. Stem combined with leaves to the right represents reading comprehension and stem combined with leaves to the left word recognition. For instance, in the range .2 to .3 one intervention caused an ES of .28 on reading comprehension and another caused an ES of .27 on word recognition. Note that many *d*s for reading comprehension concentrate between .3 and .4 and *d*s for word recognition between .5 and .6.

Quality of Meta-Analyses

Most syntheses of research satisfy the criterion that effect sizes across comparisons are independent (68%). Intercoder reliability for coding the set of studies on these methodological characteristics and hereafter discussed measures was satisfactory. Reliabilities of moderator variables are not always reported (36%) and neither do meta-analysts always make an estimate of a publication bias (21%). To prevent independence of effect sizes various strategies were used. Some adjusted sample size for significance tests so that a single subject's data did not count more than once (e.g., Ritter et al., 2009). In other studies a combined effect is estimated, and subsequent contrasts between two or more kinds of interventions are not tested (e.g., Bus & van IJzendoorn, 1999). Some studies ignore the problem and use the same control group more than once (e.g., Ehri, Nunes, Willows, et al., 2001).

In most cases, *Q* statistics are reported (70%), but a majority of studies applied a fixed model even though the populations did not involve a common effect size estimate as is indicated by the tests of homogeneity (e.g., Bus et al., 1995; Elbaum, Vaughn, Hughes, & Moody, 2000) or did not report which model was used at all (e.g., Sencibaugh, 2007; Therrien, 2004). Sometimes authors may draw strong conclusions and bold implications for practice from a combined effect size even though the point estimate is not representative of the central tendency in the total set of studies. In that case, conclusions are at least premature. Large variation in effect sizes requires a random-effects model, which implies a broader confidence interval and a higher chance that the effect size is not significantly different from zero. This scenario, however, not always holds, as can be illustrated for the book-reading study. We reanalyzed the data of the book-reading meta-analysis with a random-effects model because the overall point estimate of effect size was not based on a homogeneous set of studies (Bus et al., 1995) and found outcomes that were very similar to those resulting from a fixed model. A point estimate of $r = .27$ for the overall effect of book reading on emergent literacy, reading achievement in school age, and language skills remains significant as is indicated by a 95% CI ranging from .21 to .32. Our meta-analytic update of parent-child book reading, covering studies between 1994 and 2008, showed almost identical random effects: $r = .34$ (*95%* CI = .26, .40) for oral language and $r = .28$ (*95%* CI = .22, .36) for emergent literacy (see chapter 2, this thesis).

Future of Reading

Academic achievement trajectories are rather stable from early childhood to adolescence: Children who are among the lower third of their class when they start formal schooling are likely to remain the lower-achieving students in high school (Entwisle, Alexander, & Olson, 2005). Ideally, offering young children a stimulating home environment prevents them from starting school already lagging behind, and thereby positively influences their academic achievement trajectory. Training parents or preschool and kindergarten teachers how to read interactively (e.g., ask story-related questions) seems a promising venue to expand children's oral language (see chapter 3 and 4) as well as knowledge about the basics of reading (see chapter 4). In conventional readers, chances at academic success may also increase when students are offered interventions that improve technical reading and reading comprehension skills as appears from our meta-analysis about meta-analyses (see chapter 5). However, the development and implementation of successful interventions takes considerable amounts of money and research efforts, and programs often stop as soon as the researcher has left the side. As the number of books that are published for children and adults keeps on increasing, and as our meta-analysis shows that mere exposure to books significantly relates to not only comprehension and technical reading and spelling skills throughout development but also to more general achievement measures such as intelligence and eligibility tests for university (see chapter 2), we wonder: What are promises and pitfalls of approaching "just" reading books as an intervention in itself?

One of the major challenges is to get age-appropriate books in the homes and hands of parents and children (e.g., Guo & Harris, 2000; Neuman & Celano, 2006). As a means, books not only offer a meaningful context for words but also communicate more general knowledge. Books often describe events, cultures, or reasoning that readers are not likely to experience in daily life. Fiction, in particular, stimulates and develops imagination by offering readers the opportunity to "try on" mental states, values, and/or life experiences of characters (Harding, 1962; Oatley, 1999; Zunshine, 2006). Furthermore, reading fiction is thought to cause catharsis – a relief of burdensome emotions (Hakemulder, 2000). Enjoying books and learning from reading, however, also depends on the match between the difficulty level of a book and a reader's ability level (Carver & Leibert, 1995; Kim & Guryan, 2010). In school, children with reading difficulties often find themselves in materials that are too difficult for them, whereas the books they may choose to read during their leisure time are way too easy and/or not interesting thematically (Spear-Swerling, Brucker, & Alfano, 2010). When children get help from their parents and/or teachers in selecting stimulating books, they may get encouraged to keep on reading independently (Allington & McGill-Franzen, 2008; Kim & White, 2008).

A fascinating question that is hardly studied in the field of reading research thus far is: How may reading books lead to a book reading routine in which proficient as well as poor readers choose to read during leisure time? First, introducing books to very young children seems to stimulate interest in the world of stories, words, and written text later on (Fletcher & Reese, 2005). As early gaps in language skills reduce children's capacity to benefit from book sharing when they are 3- to 5-years old, nowadays more practitioners believe that a very early start with book reading (i.e., in the first year of life) may be of vital importance (Bus, Leseman, & Neuman, in press). Exemplary examples that promote such an early start are the American project "*Reach Out and Read* (ROR)" (Needlman & Silerstein, 2004; Needlman, Toker, Dreyer, Klass, & Mendelsohn, 2005) as well as the UK project *BookStart* (e.g., Hall, 2001) that is adopted by several other European countries among which the Netherlands. However, we have to acknowledge that reading to children in their first two years of life is especially demanding for caregivers as it requires them to react promptly and adequately to signals of distress of their young child (e.g., Bus & van IJzendoorn, 1997; Bus, Belsky, van IJzendoorn, & Crnk, 1997; DeLoache & DeMendoza, 1987). We expect therefore that the success of early interventions is uncertain when parents receive a package of books for their baby or infant without further support. Especially when parents do not use books as a source of entertainment themselves, they may not succeed to pass on pleasure in reading to their children and, hence, their children may be less likely to enjoy reading when they are able to read themselves (Bus, Leseman, & Keultjes, 2000).

Second, interactive reading has a positive effect on comprehension and thereby most likely also on reading interest. Therefore, numerous early intervention studies promote interactive book reading. However, our findings reveal several drawbacks for such an approach. On the one hand, children who were at risk for language and literacy impairments did not benefit from the interventions when their parents read to them interactively (see chapter 3). On the other hand, in a set of classroom-based studies with predominantly children at risk, most explicit effect sizes were found for experiments that were highly controlled and executed by researchers, whereas teachers who delivered interventions seemed to have difficulty with fostering the same growth in young children's language and literacy skills as researchers (see chapter 4). We hypothesized that teachers are not successful in incorporating and internalizing dialogic strategies because they are less well educated in theories of how children can benefit most from exposure to books (Dickinson & Sprague, 2001). To stimulate the use of an interactive book reading style at home and at school, it may be critical that the social component of the implementation process is emphasized more with several opportunities for feedback and positive reinforcement next to the training in more technical aspects such as the theory behind the intervention (e.g., Shernoff & Kratochwill, 2007).

Third, when children get older, their preferences for leisure-time activities may be determined increasingly less by their home environment and may depend increasingly more on their reading abilities and their attitudes and motivation towards reading books (Harlaar, Dale, & Plomin, 2007; Petrill, Deater-Deckard, Schatschneider, & Davis, 2005). However, promotion of book reading seems to remain important in all age groups. A promising finding in the first meta-analysis in this thesis is that all readers can benefit from reading books, regardless of their reading abilities. Further research is required to highlight why some poor readers are inclined to read books despite of their reading problems and get better in reading as a result of print exposure while other poor readers hardly read and increasingly fall behind.

In sum, whether children develop a book reading routine depends not only on the presence of books in children's homes and/or classrooms. Young children need help to understand stories and some parents and teachers seem to need training to get the most out of book reading. When children have become conventional readers, leisure time reading seems to make a huge difference for their cognitive development. A growing number of alternative activities in the present computer era may decrease the amount of reading time and, consequently, cause negative effects on students' language proficiency, reading skills, and broader cognitive development. Apart from the question how book reading behavior can be promoted for both proficient and poor readers, there are other fascinating questions to be studied. One of the most fascinating issues for future research might be how sharing books in infancy turns into choosing to read as a leisure-time activity in adolescence and adulthood. The answer to this question can have far-reaching consequences for education and educational policy and need to be high on the current research agenda.

Appendix 5.1

Focal Questions in Meta-Analyses in the Domain of Reading

Book reading

- Is there a relation between parent-preschooler book reading and emergent and conventional reading? (Bus et al., 1995; chapter 2, this thesis)
- Does book reading in schools affect oral language and reading skills? (Blok, 1999)
- Does dialogic reading intensify the effects of parent-child picture storybook sharing? (chapter 3, this thesis)
- Does trained interactive teacher behavior as a part of book reading improve young children's language and print-related skills? (chapter 4, this thesis)
- Do shared-reading interventions impact young children's early literacy skills? (NELP, 2008, Chapter #4)

Phonemic awareness instruction

- Does phonemic awareness training affect learning-to-read processes in a positive and substantial way, and are programs combining phonemic awareness training with letters and words more effective? (Bus & van IJzendoorn, 1999)
- Is phonemic awareness instruction effective in helping children learn to read? If so, under what circumstances and for what children? (Ehri, Nunes, Willows, et al., 2001)

Preschool intervention

- Do preschool intervention programs cause a positive effect on reading achievement? (Goldring & Pressbrey, 1986; NELP, 2008, Chapter #6)
- What are the effects of preschool programs on children's intellectual, socio-emotional, and language abilities? (Leseman et al., 1998, 1999)
- Are vocabulary interventions effective for teaching words to preschool and kindergarten children, and can vocabulary training narrow the achievement gap? (Marulis & Neuman, 2010)
- Are emergent literacy interventions with a family-component applicable for low-income, ethnic-minority, or linguistically-diverse preschool children? (Manz et al., in press)

Beginning reading methods

- Are whole-language or language experience approaches more effective than basal readers? (Stahl & Miller, 1989)
- Is whole-language instruction effective compared with basal instruction for kindergarten to third-grade students with low socioeconomic status? (Jeynes & Littell, 2000)
- Does systematic phonics instruction help children learn to read more effectively than nonsystematic phonics instruction or instruction teaching no phonics (i.e., language activities)? (Camilli, Vargas, & Yurecko, 2003; Camilli, Wolfe, & Smith, 2006; Ehri, Nunes, Stahl, & Willows, 2001; Hammill & Swanson, 2006; Stuebing, Barth, Cirino, Francis, & Fletcher, 2009)
- What are the effects of alphabet training (i.e., letter name and/or letter sound instruction with or without phonemic awareness instruction) in preschool, kindergarten, and early elementary school on the acquisition of emergent literacy skills? (Piasta & Wagner, 2009)

Reading comprehension instruction

- Does vocabulary instruction affect reading comprehension? (Stahl & Fairbanks, 1986)
- Does sentence-combining promote reading comprehension? (Fusaro, 1992)
- Does instruction in question asking affect reading comprehension? (Rosenshine, Meister, & Chapman, 1996)
- Which forms of comprehension instruction improve reading comprehension? (NRP, 2000)
- How effective is repeated reading on comprehension and what are essential instructional components? (Therrien, 2004)
- Does enhancing students' reading engagement increase reading comprehension? (Guthrie et al., 2007)
- What is the role of classroom discussion on students' text comprehension? (Murphy et al., 2009)

Acquiring vocabulary through reading

- Does instruction in deriving meaning from context improve skills to derive meaning from context? (Fukkink & De Glopper, 1998)
- Do children incidentally derive new words from texts? (Swanborn & de Glopper, 1999)

Effects of multimedia

- Does the Lightspan program (computer games to improve school-based achievement) improve reading comprehension, reading vocabulary, sounds/letters, and word reading? (Blanchard & Stock, 1999)
- How effective are computer-assisted instruction programs in the phase of beginning reading? (Blok et al., 2002)
- What is the effectiveness of information and communication technology on the teaching of spelling? (Torgerson & Elbourne, 2002)
- How effective is the use of technology (e.g., electronic books) in language education and language learning? (Zhao, 2003; Zucker et al., 2009)
- What is the effect of using glosses (e.g., level of instruction, text type) in multimedia learning environments for enhancing L2 reading comprehension? (Abraham, 2008; Taylor, 2006)
- What is the impact of digital tools on the reading performance of middle school students? (Moran et al., 2008)

Other aspects of reading instruction

- Does some form of guided oral reading stimulate reading achievement? (NRP, 2000)
- What is the impact of summer school programs (i.e., remedial, acceleration) on students' reading skills? (Cooper et al., 2000)
- Do cognitive paradigms targeting domain-specific learning activities improve effectiveness of reading instruction? (Seidel & Shavelson, 2007)
- What is the effect of morphological instruction in elementary school on reading and spelling development? (Bowers, Kirby, & Deacon, in press)

Bilingual children

- Does learning to read in the native language promote reading achievement in the second language? (Greene, 1997; Rolstad et al., 2005; Slavin & Cheung, 2005; Willig, 1985)
- Is bilingualism related to cognitive variables such as literacy and metalinguistic awareness (Adesope, Lavin, Thompson, & Ungerleider, 2010)

Instruction of children with reading disabilities

- What is the overall effectiveness of sight word teaching for individuals with moderate and severe disabilities? (Browder & Xin, 1998)
- Does direct instruction yield higher effect sizes than strategy instruction in groups with learning disabilities? (Swanson, 1999; Swanson & Hoskyn, 1998)
- Do studies using strategy instruction or direct instruction yield higher effect size estimates than studies using competing models? (Swanson & Sachse-Lee, 2000)
- Do instructional components predict positive outcomes for adolescents with learning disabilities on measures of higher order processing? (Swanson, 2001)
- How effective is the Reading Recovery program for low-performing first-grade students? (D'Agostino & Murphy, 2004)
- Does treatment to improve expressive or receptive phonology, syntax, or vocabulary affect children with primary developmental speech and language disorders? (Law, Garrett, & Nye, 2004)
- What is the supplemental effect of out-of-school programs on reading achievement of at-risk students from kindergarten to high school? (Lauer et al., 2006)
- Do metacognitive strategies improve the reading comprehension levels of students with learning disabilities? (Sencibaugh, 2007)
- How do interventions targeting decoding, fluency, vocabulary, and comprehension influence comprehension outcomes for secondary students with reading difficulties? (Edmonds et al., 2009)

Effects of grouping and tutoring

- Does one-to-one tutoring on reading promote reading skills? (Elbaum, Vaughn, Hughes, & Moody, 2000)
- Is effect size of reading instruction related to grouping format (e.g., pairing, small groups)? (Elbaum et al., 1999; Elbaum, Vaughn, Hughes, Moody, et al., 2000)
- Is parental involvement related to children's academic achievement (i.e., reading)? (Fan & Chen, 2001; Jeynes, 2002, 2005; NELP, 2008, Chapter #5; Sénéchal & Young, 2008)
- Do volunteer tutoring programs in elementary and middle school improve reading skills? (Ritter et al., 2009; Torgerson et al., 2002)

Effects of school organization

- Do second graders who have received 2 years of instruction in smaller classes score significantly higher in reading than second graders in larger classes? (McGiverin et al., 1989)
- Does reading achievement decline over summer holiday? (Cooper et al., 1996)
- Does homework improve academic achievement (i.e., reading)? (Cooper, Robinson, & Patall, 2006)

Processes explaining reading (dis)abilities

- Are auditory perception skills related to reading? (Kavale, 1980, 1981)
- Is visual perception an important correlate of reading achievement? (Kavale, 1982)
- Which of six variables (language, sensory skills, behavioral-emotional, soft neurological, IQ, and teacher ratings) provide the best early prediction of later reading difficulties? (Horn & Packard, 1985)
- Do dyslexic readers and normal readers differ in terms of phonological skill despite equivalent word-recognition abilities? (Herrmann, Matyas, & Pratt, 2006; van IJzendoorn & Bus, 1994)
- Do measures that tax the processing as well as the storage resources of working memory predict reading comprehension better than measures that tax only the storage resources? (Daneman & Merikle, 1996)
- Is a regularity effect also present in a group with learning disabilities? (Metsala, Stanovich, & Brown, 1998)
- Do children with learning disabilities differ from normal-achieving children in immediate memory performance, and does this difference continue? (O'Shaughnessy & Swanson, 1998)
- Do underachieving students with and without a learning disabilities label differ in reading performance? (Fuchs, Fuchs, Mathes, & Lipsey, 2000)
- Do children with reading disabilities and low achievers share a common deficit in phonological processing, memory, and visual-spatial reasoning? (Hoskyn & Swanson, 2000)
- What is the relative importance of auditory and visual perception in predicting reading achievement? (Kavale & Forness, 2000)
- Is it valid to use IQ discrepancy for the classification of reading disabilities? (Steubing et al., 2002)
- Which brain areas are consistently activated during aloud single word-reading tasks? (Turkeltaub, Eden, Jones, & Zeffiro, 2002)
- Are rapid naming and phonological awareness as strong predictors of word reading as related reading abilities? (Swanson, Trainin, Necoechea, & Hammill, 2003)
- What is the influence of school mobility in the United States on reading achievement in the elementary grades? (Mehana & Reynolds, 2004)
- Can the relative variability of psychophysical performance in dyslexic readers compared with normal readers be attributed to general nonsensory difficulties? (Roach, Edwards, & Hogben, 2004)
- Does sampling affect studies linking genes to complex phenotypes such as reading ability/disability and related componential processes? (Grigorenko, 2005)
- What are the patterns of convergence in neuroanatomical circuits underlying phonological processing in reading alphabetic words and logographic characters? (Tan, Laird, Li, & Fox, 2005)
- Are gender differences present in reading achievement, and do these change with age? (Lietz, 2006; Lynn & Mikk, 2009)
- What is the magnitude and consistency of balance difficulties in the dyslexia population and which sampling or stimulus characteristics modulate this effect? (Rochelle & Talcott, 2006)

- Do children with and without specific language impairments show performance differences in nonword repetition? (Graf Estes, Evans, & Else-Quest, 2007)
- What are the links between school entry skills or school readiness and later school reading achievement? (Duncan et al., 2007; La Paro & Pianta, 2000)
- What is the role of executive functioning measures (e.g., task modality) in distinguishing between performance of children with and without reading difficulties? (Booth, Boyle, & Kelly, 2010; Carretti, Borella, Cornoldi, & De Beni, 2009)
- Do children with type I diabetes perform lower than children without diabetes on a variety of cognitive domains including reading and writing? (Naguib, Kulinskaya, Lomax, & Garralda, 2009)
- What is the association between the Oral Reading measure of Curriculum-Based Measurement and other standardized measures of reading achievement for students in grade 1 to 6? (Reschly, Busch, Betts, Deno, & Long, 2009)
- What functional abnormalities in the brain are consistently associated with dyslexia? (Richlan, Kronbichler, & Wimmer, 2009)
- To what extent and in what manner do adults with reading disabilities differ from adults without reading disabilities on measures assumed to relate to overall reading competence? (Swanson & Hsieh, 2009)
- Does the magnitude of cognitive processing differences (e.g., reading, oral language) between students with specific learning disabilities and typically achieving peers justify inclusion in classification of SLD? (Johnson, Humphrey, Mellard, Woods, & Swanson, 2010)

References

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning, 21,* 199-226.

Adams, M. J. (1990). *Beginning to read. Thinking and learning about print*. Cambridge, MA: MIT Press.

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80,* 207-245.

Allington, R., & McGill-Franzen, A. (2008). Got books? *Educational Leadership. 65*, 20-23.

Blanchard, J., & Stock, W. (1999). Meta-analysis of research on a multimedia elementary school curriculum using personal and video-game computers. *Perceptual and Motor Skills, 88,* 329–336.

Blok, H. (1999). Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning, 49,* 343–371.

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72,* 101-130.

Booth, J. N., Boyle, J. M. E., & Kelly, S. W. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *British Journal of Developmental Psychology, 28,* 133-176.

Bowers, P., N., Kirby, J. R., & Deacon, S. H. (in press). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research.*

Browder, D. M., & Xin, Y. P. (1998). A meta-analysis and review of sight word research and its implications for teaching functional reading to individuals with moderate and severe disabilities. *Journal of Special Education, 32,* 130–153.

Bus, A. G., & van IJzendoorn, M. H. (1997). Affective dimension of mother-infant picturebook reading. *Journal of School Psychology, 35,* 47-60.

Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91,* 403–414.

Bus, A. G., Leseman, P. P., & Keultjes, P. (2000). Joint book reading across cultures: A comparison of Surinamese-Dutch, Turkish-Dutch, and Dutch parent-child dyads. *Journal of Literacy Research, 32,* 53-76.

Bus, A.G., Leseman, P.M., & Neuman, S.B. (in press). Methods for Preventing Early Academic Difficulties. In S. Graham, A.G. Bus, S. Major, & L. Swanson (Eds.), *APA Handbook of Educational Psychology, Volume 3* (Application of Educational Psychology to Learning and Teaching).

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1-21.

Bus, A. G., Belskey, J., van IJzendoorn, M. H., & Crnic, K. (1997). Attachment and bookreading patterns: A study of mothers, fathers, and their toddlers. *Early Childhood Research Quarterly, 12,* 81-98.

Bushman, B. J., & Wells, G. L. (2001). Narrative impressions of literature: The availability bias and the corrective properties of meta-analytic approaches. *Personality and Social Psychology Bulletin, 27,* 1123-1130.

Camilli, G., Vargas, S., & Yurecko, M. (2003). Teaching children to read: The fragile link between science and federal education policy. *Education Policy Analysis Archives, 11*. Retrieved November 11, 2009, from http://epaa.asu.edu/epaa/v11n15/.

Camilli, G., Wolfe, P. M., & Smith, M. L. (2006). Meta-analysis and reading policy: Perspectives on teaching children to read. *Elementary School Journal, 107,* 27-36.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19,* 246-251.

Carver, R. P., & Leibert, R. E. (1995). The effect of reading library books at different levels of diffuculty upon gains in reading ability. *Reading Research Quarterly, 30*, 26-48.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65,* 1-127.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd edition). New York: Russell Sage Foundation.

Cooper, H., Nye, B., Charlton, K., & Lindsay, J. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66,* 227–268.

Cooper, H., J. C., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987-2003. *Review of Educational Research, 76,* 1-62.

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin, 87,* 442–449.

Cunningham, J. W. (2001). The National Reading Panel report. *Reading Research Quarterly, 36*, 326–335.

D'Agostino, J. V., & Murphy, J. A. (2004). A meta-analysis of reading recovery in United States schools. *Educational Evaluation and Policy Analysis, 26,* 23-38.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review, 3,* 422–433.

Dickinson, D. K., & Sprague, K. E. (2001). The nature and impact of early childhood care environments on the language and early literacy development of children from low-income families. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of Early Literacy Research* (Vol. 1, pp. 263-280). New York: The Guilford Press.

DeLoache, J. S., & Demendoza, O. A. P. (1987). Joint picturebook interactions of mothers and 1-year-old children. *British Journal of Developmental Psychology, 5,* 111-123.

Duncan, G. J., Claessens, A., Huston, A. C., Pagani, L. S., Engel, M., Sexton, H. et al. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428-1446.

Eagly, A. H., & Wood, W. (1994). Tying research synthesis to substantive issues. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 485–502). New York: Russell Sage.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K. et al. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79,* 262-300.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393–447.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36,* 250–287.

Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (1999). Grouping practices and reading outcomes for students with disabilities. *Exceptional Children, 65,* 399–415.

Elbaum, B., Vaughn, S., Hughes, M. T., Moody, S. W., & Schumm, J. S. (2000). How reading outcomes of students with disabilities are related to instructional grouping formats: A meta-analytic review. In R. M. Gersten & E. P. Schiller (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues. The LEA series on special education and disability* (pp. 105–135). Mahwah, NJ: Erlbaum.

Elbaum, B., Vaughn, S., Hughes, M. T., & Watson-Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92,* 605–619.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First grade and educational attainment by age 22: A new story. *American Journal of Sociology, 110,* 1458-1502.

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13,* 1-22.

Fletcher, K. L., & Reese, E. (2005). Picture book reading with young children: A conceptual framework. *Developmental Review, 25,* 64-103.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Lipsey, M. W. (2000). Reading differences between low-achieving students with and without learning disabilities: A meta-analysis. In R. M. Gersten & E. P. Schiller (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp. 81–104). Mahwah, NJ: Erlbaum.

Fukkink, R. G., & de Glopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research, 68,* 450–469.

Fusaro, J. A. (1992). Meta-analysis of the effect of sentence-combining on reading comprehension when the criterion measure is the test of reading comprehension. *Perceptual and Motor Skills, 74,* 331–333.

Garfield, E. (1979). *Citation indexing. Its theory and application in science, technology, and humanities.* New York: Wiley.

Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Research, 5*, 3–8.

Goldring, E. B., & Pressbrey, L. S. (1986). Evaluating preschool programs: A meta-analytic approach. *Educational Evaluation and Policy Analysis, 8,* 179–188.

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50,* 177-195.

Greene, J. P. (1997). A meta-analysis of the Rossell and Baker review of bilingual education research. *Bilingual Research Journal, 21,* 103-122.

Grigorenko, E. L. (2005). A conservative meta-analysis of linkage and linkage-association studies of developmental dyslexia. *Scientific Studies of Reading, 9,* 285-316.

Guo, G., & Harris, K. M. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography, 37,* 431-447.

Guthrie, J. T., McRae, A., & Klauda, S. L. (2007). Contributions of concept-oriented reading instruction to knowledge about interventions for motivations in reading. *Educational Psychologist, 42,* 237-250.

Hakemulder, J. (2000). *The moral laboratory: Experiments examining the effects of reading literature on social perception and moral self-concept.* Amsterdam, The Netherlands: John Benjamins B.V.

Hammill, D. D., & Swanson, H. L. (2006). The National Reading Panel's meta-analysis of phonics instruction: Another point of view. *The Elementary School Journal, 107,* 17-26.

Harding, D. W. (1962). Psychological processes in the reading of fiction. *British Journal of Aesthetics, 2,* 133-147.

Hedges, L. V. (2008). What are effect sizes and why do we need them? *Child Development Perspectives, 2,* 167-171.

Herrmann, J. A., Matyas, T., & Pratt, C. (2006). Meta-analysis of the nonword reading deficit in specific reading disorder. *Dyslexia, 12,* 195-221.

Horn, W. F., & Packard, T. (1985). Early identification of learning problems: A meta-analysis. *Journal of Educational Psychology*, *77*, 597–607.

Hoskyn, M., & Swanson, H. L. (2000). Cognitive processing of low achievers and children with reading disabilities: A selective meta-analytic review of the published literature. *School Psychology Review, 29,* 102-119.

Jeynes, W. H. (2002). A meta-analysis: The effects of parental involvement on minority's children's academic achievement. *Education and Urban Society, 35,* 202-218.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education, 40,* 237-269.

Jeynes, W. (in press). The relationship between Bible literacy and behavioral and academic outcomes in urban areas: A meta-analysis. *Education and Urban Society.*

Jeynes, W. H., & Littell, S. W. (2000). A meta-analysis of studies examining the effect of whole language instruction on the literacy of low-SES students. *The Elementary School Journal*, *101*, 21–33.

Johnson, E. S., Humphrey, M., Mellard, D. F., Woods, K., & Swanson, H. L. (2010). Cognitive processing deficits and students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly, 33,* 3-18.

Hall, E. (2001). Babies, books, and 'impact': Problems and possibilities in the evaluation of a Bookstart project. *Educational Review, 53,* 57-64.

Harlaar, N., Dale, P. S., & Plomin, R. (2007). Reading exposure: A (largely) environmental risk factor with environmentally-mediated effects on reading performance in the primary school years. *Journal of Child Psychology and Psychiatry, 48,* 1192-1199.

Kavale, K. A. (1980). Auditory-visual integration and its relationship to reading achievement: A meta-analysis. *Perceptual and Motor Skills, 51,* 947–955.

Kavale, K. A. (1981). The relationship between auditory perceptual skills and reading ability: A meta-analysis. *Journal of Learning Disabilities, 14,* 539–546.

Kavale, K. A. (1982). Meta-analysis of the relationship between visual perceptual skills and reading achievement. *Journal of Learning Disabilities, 15,* 42–51.

Kavale, K. A., & Forness, S. R. (2000). Auditory and visual perception processes and reading ability: A quantitative reanalysis and historical reinterpretation. *Learning Disability Quarterly, 23,* 253–270.

Kim, J. S., & Guryan, J. (2010). The efficacy of voluntary summer book reading intervention for low-income Latino children from language minority families. *Journal of Educational Psychology, 102,* 20-31.

Kim, J. S., & White, T. G. (2008). Scaffolding voluntary summer reading for children in grades 3 to 5: An experimental study. *Scientific Studies of Reading, 12,* 1-23.

La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research, 70,* 443-484.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76,* 275-313.

Law, J., Garrett, Z., & Nye, C. (2004). The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 47,* 924-943.

Leseman, P. P. M., Otter, M. E., Blok, H., & Dekkers, P. (1998). Effecten van voor- en vroegschoolse educatieve centrumprogramma's. Een meta-analyse van studies gepubliceerd tussen 1985 en 1996 [Effects of preschool intervention programs; A meta-analysis of studies published between 1985 and 1996]. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs, 14,* 134-154.

Leseman, P. P. M., Otter, M. E., Blok, H., & Dekkers, P. (1999). Effecten van voorschoolse educatieve centrumprogramma's. Een aanvullende meta-analyse van studies gepubliceerd tussen 1985 en 1996 [Effects of preschool intervention programs; An additional meta-analysis of studies published between 1985 and 1996]. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs, 15,* 28-37.

Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies in Educational Evaluation, 32,* 317-344.

Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *Trames, 13,* 3-13.

Lytton, H. (1994). Replication and meta-analysis: The story of a meta-analysis of parents' socialization practices. In R. van der Veer, M. H. van IJzendoorn & J. Valsiner (Eds.), *Reconstructing the mind: Replicability in research on human development* (pp. 117–150). Norwood, NJ: Ablex.

Manz, P. H., Hughes, C., Barnabas, E., Bracaliello, C., & Ginsburg-Block, M. (in press). A descriptive review and meta-analysis of family-based emergent literacy interventions: To what extent is the research applicable to low-income, ethnic-minority or linguistically-diverse young children? *Early Childhood Research Quarterly*.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research, 80,* 300-335.

McGiverin, J., Gilman, D., & Tillitski, C. (1989). A meta-analysis of the relation between class size and achievement. *Elementary School Journal, 90,* 47–56.

Mehana, M., & Reynolds, A. J. (2004). School mobility and achievement: A meta-analysis. *Children and Youth Services Review, 26,* 93-119.

Metsala, J. L., Stanovich, K. E., & Brown, G. D. A. (1998). Regularity effects and the phonological deficit model of reading disabilities: A meta-analytic review. *Journal of Educational Psychology, 90,* 279–293.

Moran, J., Ferdig, R. E., Pearson, P. D., Wardrop, J., & Blomeyer, R. L. (2008). Technology and reading performance in the middle-school grades: A meta-analysis with recommendations for policy and practice. *Journal of Literacy Research, 40,* 6-58.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101,* 740-764.

Naguib, J. M., Kulinskaya, E., Lomax, C. L., & Garralda, M. E. (2009). Neuro-cognitive performance in children with type 1 diabetes – A meta-analysis. *Journal of Pediatric Psychology, 34,* 271-282

National Early Literacy Panel (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Rockville, MD: NICHD Clearinghouse.

Needlman, R., & Silverstein, M. (2004). Pediatric interventions to support reading aloud: How good is the evidence? *Developmental and Behavioral Pediatrics, 25,* 352-363.

Needlman, R., Toker, K. H., Dreyer, B. P., Klass, P., & Mendelsohn, A. L. (2005). Effectiveness of a primary care intervention to support reading aloud: A multicenter evaluation. *Ambulatory Pediatrics, 5,* 209-215.

Neuman, S. B., & Celano, D. (2006). The knowledge gap: Implications of leveling the playing field for low-income and middle-income children. *Reading Research Quarterly, 41,* 176-201.

Oatley, K. (1999). Why fiction may be twice as true as fact: Fiction as cognitive and emotional simulation. *Review of General Psychology, 3,* 101-117.

O'Shaughnessy, T. E., & Swanson, H. L. (1998). Do immediate memory deficits in students with learning disabilities in reading reflect a developmental lag or deficit?: A selective meta-analysis of the literature. *Learning Disability Quarterly, 21,* 123–148.

Pearson, K. (1904). Report on certain enteric fever inoculation statistics. *British Medical Journal, 3*, 1243–1246.

Piasta, S. B., & Wagner, R. K. (2009). Developing early literacy skills: A meta-analysis of alphabet learning and instruction. *Reading Research Quarterly, 45,* 8-38.

Petrill, S. A., Deater-Deckard, K., Schatschneider, C., & Davis, C. (2005). Measured environmental influences on early reading: Evidence from an adoption study. *Scientific Studies of Reading, 9,* 237-259.

Pressley, M., & Harris, K. R. (1994). Increasing the quality of educational intervention research. *Educational Psychology Review*, *6*, 191–208.

Reschley, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-Based Measurement Oral Reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47,* 427-469.

Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping, 30,* 3299-3308.

Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research, 79,* 3-38.

Roach, N. W., Edwards, V. T., & Hogben, J. H. (2004). The tale is in the tail: An alternative hypothesis for psychophysical performance variability in dyslexia. *Perception, 33,* 817-830.

Rochelle, K. S. H., & Talcott, J. B. (2006). Impaired balance in developmental dyslexia? A meta-analysis of the contending evidence. *Journal of Child Psychology and Psychiatry, 47,* 1159-1166.

Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy, 19,* 572-594.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, *66*, 181–221.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77,* 454-499.

Sencibaugh, J. M. (2007). Meta-analysis of reading comprehension interventions for students with learning disabilities: Strategies and implications. *Reading Improvement, 44,* 6-22.

Sénéchal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. *Review of Educational Research, 78,* 880-907.

Shernoff, E. S., & Kratochwill, T. R. (2007). Transporting an evidence-based classroom management program for preschoolers with disruptive behavior problems to a school: An analysis of implementation, outcomes, and contextual variables. *School Psychology Quarterly, 22,* 449-472.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, *31*, 15–21.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75,* 247-284.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Spear-Swerling, L., Brucker, P. O., & Alfano, M. P. (2010). Relationships between sixth-graders' reading comprehension and two different measures of print exposure. *Reading & Writing, 23,* 73-96.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56,* 72–110.

Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research, 59*, 87–116.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-407.

Stanovich, K. E. (2000). *Progress in understanding reading; Scientific foundations and new frontiers.* New York: The Guilford Press.

Stuebing, K. K., Barth, A. E., Molfese, P. J., Weiss, B., & Fletcher, J. M. (2009). IQ is not strongly related to response to reading instruction: A meta-analytic interpretation. *Exceptional Children, 76,* 31-51.

Stuebing, K. K., Fletcher, J. M., Ledoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39,* 469-518.

Sutton, A. J., & Higgins, J. P. T. (2008). Recent developments in meta-analysis. *Statistics in Medicine, 27,* 625-650.

Swanborn, M. S. L., & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69,* 261–285.

Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis in intervention outcomes. *Journal of Learning Disabilities, 32,* 504–532.

Swanson, H. L. (2001). Research on interventions for adolescents with learning disabilities: A meta-analysis of outcomes related to higher-order processing. *Elementary School Journal, 101,* 331–348.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research, 68*, 277–321.

Swanson, H. L., & Hsieh, C.-J. (2009). Reading disabilities in adults: A selective meta-analysis of the literature. *Review of Educational Research, 79,* 1362-1390.

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities, 33,* 114–136.

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73,* 407-440.

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping, 25,* 83-91.

Taylor, A. (2006). The effects of CALL versus traditional L1 glosses on L2 reading comprehension. *CALICO Journal, 23,* 309-318.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and Special Education, 25,* 252-261.

Torgerson, C. J., & Elbourne, D. (2002). A systematic review and meta-analysis of the effectiveness of information and communication technology (ICT) on the teaching of spelling. *Journal of Research in Reading, 25,* 129-143.

Torgerson, C. J., King, S. E., & Sowden, A. J. (2002). Do volunteers in schools help children learn to read? A systematic review of randomised controlled trials. *Educational Studies, 28,* 433-444.

Torgerson, C. J., Porthouse, J., & Brooks, G. (2003). A systematic review and meta-analysis of randomised controlled trials evaluating interventions in adult literacy and numeracy. *Journal of Research in Reading, 26,* 234-255.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage, 16,* 765-780.

Van IJzendoorn, M. H. (1994). Process model of replication studies: On the relations between different types of replication. In R. van der Veer, M. H. van IJzendoorn & J. Valsiner (Eds.), *Reconstructing the mind: Replicability in research on human development* (pp. 57–70). Norwood, NJ: Ablex.

Van IJzendoorn, M. H., & Bus, A. G. (1994). Meta-analytic confirmation of the nonword reading deficit in developmental dyslexia. *Reading Research Quarterly, 29,* 266–275.

Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research, 55,* 269–317.

Zao, Y. (2003). Recent developments in technology and language learning: A literature review and meta-analysis. *CALICO Journal, 21,* 7-27.

Zucker, T. A., Moody, A. K., & McKenna, M. C. (2009). The effects of electronic books on pre-kindergarten-to-grade-5 students' literacy and language outcomes: A research synthesis. *Journal of Educational Computing Research, 40,* 47-87.

Zunshine, L. (2006). Why we read fiction: Theory of mind and the novel. In J. Phelan & P. J. Rabinowitz (Eds.), *Theory and interpretation of narrative.* Ohio: Thomson-Shore.

Summary

There is a widely held belief that reading (story)books makes us smarter and helps promote success in life. Does scientific evidence support this notion? The three meta-analyses in this thesis comprise 146 studies between 1988 and 2010 (N = 10,308 participants) that addressed the role of book reading in language and reading development from infancy to early adulthood. For the group of pre-conventional readers, the effect of interventions at home and at school that improve the quality of shared book reading such as dialogic reading programs were also examined.

Before formal reading instruction starts, young children already form basic concepts about the connections between spoken and written words, which eventually leads to the ability to read and spell words fluently and accurately. As the ultimate goal of reading is reading for understanding, children's reading proficiency gets increasingly less determined by technical reading skills and gets increasingly more dependent on sophisticated vocabulary, background knowledge, and intelligence. Because books are an important means to get exposed to a variety of word meanings and word forms in relevant contexts, this thesis aimed to quantify how reading narrative texts (e.g., storybooks, novels, magazines) is related to indicators of reading comprehension and technical reading and spelling skills across development. It was hypothesized that reading books is both a consequence of reading proficiency and a contributor to further reading development: Because more skilled readers are more likely to enjoy books, they will choose to read more frequently which, in turn, will improve knowledge of word forms and semantics and enhance vocabulary size and text comprehension abilities.

A meta-analysis is a powerful tool to integrate, standardize, and systematically summarize findings of studies with comparable measures, interventions, and/or outcome domains. Effect sizes, quantitative indexes of relations among variables, are used to compare and communicate the strength of the summarized research findings. The first meta-analysis on shared book reading (Bus, Van IJzendoorn, & Pellegrini, 1995) showed moderate effect sizes for oral language and basic reading skills, which indicates that it certainly can make a tremendous difference in the lives of young children whether or not they are read to by their parent. Because meta-analyses are based on numerous decisions about collecting, coding, and analyzing the research base that far and because new interventions and measures continue to be developed and tested in different groups of children across countries, new meta-analyses are needed to replicate and extend earlier findings and to make up-to-date recommendations to the field.

The first meta-analysis in this thesis indicated that leisure-time reading activities can be considered as a driving force in shaping language and literacy. In preschool and kindergarten, grade 1 to 12, and college and university, the association between leisure-time reading activities and age-appropriate

measures in the domains of reading comprehension and technical reading and spelling skills were moderate to strong. As expected and in line with a model of reciprocal causation, leisure-time reading became increasingly more important for oral language, technical reading, and intelligence with each year of education. Impressively, exposure to books explained 12% of the variance in oral language skills in preschool and kindergarten, 13% in primary school, 19% in middle school, 30% in high school, and 34% in college and university. Furthermore, leisure time reading seemed especially important for low-ability readers: When they have experience with books at home, low-ability readers have more opportunities to practice basic reading skills, and consequently, become more accurate and fluent in text reading than their low(er)-ability peers who read less. Overall, the first meta-analysis suggested that reading routines that are part of children's and students' leisure-time activities offer substantial advantages for the development of reading proficiency and academic success.

For pre-conventional readers, books cannot be a means to stimulate language and basic reading skills as long as children do not receive intensive support from adults to remain attentive, to discover exciting parts of a story, and to understand unfamiliar words or difficult phrases. In a stimulation package called "Dialogic Reading", caregivers are trained to stimulate active involvement by eliciting verbal responses to the story with the help of open-ended questions about pictured materials and by providing informative feedback on child responses. The second meta-analysis demonstrated that enhancing the dialogue between parent and child indeed strengthened the effects of book reading. Parents who read dialogically enlarged their children's vocabularies significantly more than control-group parents who shared books as they were used to. Strikingly, two subgroups did not appear to benefit from the intervention: The oral language skills of 5- to 6-year-old kindergarten children and children at risk for language and literacy impairments hardly improved. On the one hand, expectations and dialogic-reading methods may have been pitched too low for kindergarten children, who may get distracted from the story content when there is too much talking. On the other hand, at-risk children who are most in need of effective language promotion were mostly from low socio-economic status homes. Their relatively low-educated parents might have experienced difficulty with incorporating the trained techniques.

One of the goals of the third meta-analysis was to test whether the literacy environment at school might be more stimulating for children at risk and/or in kindergarten classrooms. Furthermore, the set of studies was large enough to test whether interactive storybook reading affected oral language as well as basic reading skills such as alphabet knowledge and phonological sensitivity. Interestingly, about 7% of the growth in oral language skills of both preschoolers and kindergartners at risk could be explained by an interactive reading intervention in the educational setting. Furthermore, kindergartners seemed to be capable to independently process and learn from printed features in storybooks during interactive reading sessions as they gained significantly more alphabet knowledge than preschool children. Apparently, at-risk children as well as older children are able to benefit from interactive reading. It does not seem, however, that a wide-scale integration of interactive reading in the regular school curriculum is warranted yet. That is, researchers appeared to be largely and significantly more effective in enhancing children's oral language skills than children's own teachers, who revealed moderate effects only when reading to whole classrooms. To enhance the effectiveness of an interactive book reading style at school, therefore, it may be critical that teachers are coached individually and receive more information about the theory behind the intervention.

In sum, the meta-analyses in this thesis supported that leisure-time reading is vital for school success and that an early start with shared book reading is important for developing the knowledge required for eventual success in reading. In fact, shared book reading may be part of a continuum of out-of-school reading experiences that facilitate children's language, reading, and spelling achievement throughout students' development. It seems, therefore, a logical step to invest in improvement of the quality of book reading to young children. However, the results of two meta-analyses testing the effects of interventions at home and at school revealed disappointing results especially for groups and settings where such an improvement in high-quality interactions with books and literacy is needed most. Dialogic Reading, a program to stimulate interactive book sharing, failed in low-educated families. In schools that were predominantly attended by children at risk, Dialogic Reading and similar interactive reading programs were least successful when they were carried out by children's own teachers. Apart from studying how both the quality and quantity of book reading can be effectively promoted for young children as well as for poor and proficient readers, future research is needed that follows children longitudinally so processes and strategies can be identified that turn sharing books in infancy into choosing to read as a leisure-time activity in adolescence and adulthood.

Samenvatting (Summary in Dutch)

Over het enorme belang van lezen bestaat nauwelijks discussie. "*Lezen is een onmisbare vaardigheid in de samenleving*," schreef minister Plasterk (OCW) in het voorjaar van 2008 aan de Tweede Kamer. Een vroeg begin met voorlezen wordt daarbij gezien als een cruciale stap voor de ontwikkeling van de kennis die nodig is om een vaardige lezer te worden. Is er wetenschappelijk bewijs voor de assumptie dat lezen onmisbaar is en ons slimmer en succesvoller maakt? De drie meta-analyses in dit proefschrift vatten 146 internationale studies samen met in totaal meer dan 10.000 kinderen en studenten waarin de rol van (voor)leesgedrag in de taal- en leesontwikkeling van zeer jonge kinderen tot jongvolwassenen centraal staat. Omdat peuters en kleuters het meest profiteren van het lezen van boeken wanneer ze gerichte hulp krijgen van hun ouders en verzorgers wordt er apart aandacht besteed aan interventieprogramma's, die tot doel hebben om de kwaliteit van de voorleessessies thuis en op school te verhogen.

Kinderen die opgroeien in een omgeving met boeken weten voordat ze naar de basisschool gaan vaak al dat er een verband is tussen geschreven en gesproken taal en dat tekst van links naar rechts gelezen wordt. Deze en andere basiskennis zijn nodig om uiteindelijk vloeiend en correct te leren lezen en spellen. Naarmate kinderen betere technische lezers worden en de moeilijkheidsgraad van teksten toeneemt, wordt hun leessucces steeds meer bepaald door hun niveau van begrijpend lezen, dat mede afhangt van de grootte en breedte van de woordenschat, algemene kennis en het intelligentieniveau. Het lezen van boeken kan een belangrijke en ontspannen manier zijn om deze taal- en leesvaardigheden op te doen en uit te breiden. Het verhaal biedt bijvoorbeeld een relevante context waaruit de betekenis van moeilijke of onbekende woorden kan worden afgeleid. In dit proefschrift is op een kwantitatieve manier geanalyseerd hoe het lezen van verhalende teksten, zoals prentenboeken, romans en tijdschriften, samenhangt met indicatoren van begrijpend lezen en technisch lezen en spellen van kinderen, jongeren en jongvolwassenen. Er werd verwacht dat het lezen van boeken zowel een consequentie is van leesvaardigheid als dat het bijdraagt aan verdere leesontwikkeling. Omdat goede lezers waarschijnlijk meer plezier beleven aan boeken zullen ze ervoor kiezen om vaker te lezen, waardoor hun technische vaardigheden toenemen, hun woordenschat uitbreidt en hun tekstbegrip groter wordt.

Een meta-analyse is een krachtige methode om resultaten van studies met vergelijkbare meetinstrumenten en/of interventies te integreren, standaardiseren en op een systematische manier samen te vatten. Om onderzoeksbevindingen te kunnen vergelijken en communiceren worden effectgroottes gebruikt – kwantitatieve indices van de relatie tussen variabelen. De eerste meta-analyse naar het effect van voorlezen op de woordenschat en de basisvaardigheden van lezen (Bus, Van IJzendoorn, & Pellegrini, 1995) vond matig sterke effecten. Deze studie ondersteunde daarmee de aanname dat het een groot verschil maakt in het leven van jonge kinderen of ze worden voorgelezen of niet. Het is echter van belang om meta-analyses regelmatig te vernieuwen en uit te breiden, omdat er veel nieuwe studies zijn verschenen en er eveneens nieuwe interventies en instrumenten zijn ontwikkeld en getest in verschillende groepen kinderen over de hele wereld.

Uit de eerste meta-analyse in dit proefschrift bleek dat vrijetijdslezen een drijvende kracht is achter geletterdheid en taalvaardigheid. Het verband tussen het lezen in de vrije tijd en uitkomstmaten als begrijpend lezen en technisch lezen en spellen was gemiddeld tot sterk voor kleuters, basisschoolleerlingen, middelbare scholieren en studenten aan hbo en de universiteit. In overeenstemming met een model van reciproque causaliteit werd het verband tussen lezen in de vrije tijd en de woordenschat, het technisch lezen en het intelligentieniveau met elk schooljaar sterker. Boeken lezen verklaarde 12% in de woordenschat van peuters en kleuters, 13% in de middenbouw, 19% in de bovenbouw van de basisschool en de eerste klassen van de middelbare school, 30% van de hogere klassen van de middelbare school en 34% op hbo- en universiteitsniveau. Zwakke lezers bleken betere basisvaardigheden, zoals kennis van het alfabet en fonologische bewustzijn, te hebben als ze aangaven te lezen in hun vrije tijd. Doordat deze leerlingen meer momenten hebben waarop ze deze vaardigheden kunnen oefenen, worden ze beter in het lezen van teksten dan zwakke lezers, die minder vaak lezen. Uit de eerste meta-analyse mag worden geconcludeerd dat het substantiële voordelen heeft voor de ontwikkeling van begrijpend en technisch lezen en voor succes op school als kinderen van jongs af aan een leesroutine weten te ontwikkelen.

Zolang kinderen nog moeten worden voorgelezen is het belangrijk dat volwassenen hun aandacht sturen en hen helpen bij het ontdekken van spannende delen van het verhaal en het begrijpen van onbekende woorden of moeilijke uitdrukkingen. In een interventie die bekend staat als *Dialogic Reading* [Interactief Voorlezen] worden ouders of verzorgers getraind om kinderen actief te betrekken bij een prentenboek. Zo leren ouders open vragen te stellen over plaatjes, karakters of gebeurtenissen in een verhaal en om onderdelen van het verhaal te koppelen aan eigen ervaringen uit het dagelijkse leven, zodat het kind aangemoedigd wordt om te reageren. De tweede meta-analyse toonde aan dat het stimuleren van de dialoog tussen ouder en kind de effecten van voorlezen kan versterken. Kinderen die interactief werden voorgelezen, hadden na afloop een significant grotere woordenschat dan controlegroepkinderen die werden voorgelezen zoals hun ouders dat gewend waren. Opvallend genoeg profiteerden twee groepen kinderen niet van de interactieve voorleesinterventie: zowel de woordenschat van 5- tot 6-jarige kleuters als de woordenschat van kinderen met een verhoogd risico op taal- en leesachterstanden bleek niet uit te breiden. Het interactieve voorlezen zou de oudere kleuters teveel kunnen afleiden van de tekst, waaruit ze de meeste nieuwe woorden leren. Voor ouders van risicogroepkinderen is het voorstelbaar dat ze het, bijvoorbeeld door onervarenheid met voorlezen, te lastig vonden om de getrainde voorleestechnieken in praktijk te brengen.

Het was een van de doelen van de derde meta-analyse in dit proefschrift om te testen of risicogroepkinderen wel konden profiteren van interactief voorlezen op de peuterspeelzaal en in de kleuterklas. Veel van de interventiestudies, die werden uitgevoerd op scholen waar vooral risicogroepen les kregen, rapporteerden naast woordenschattaken ook testen van basisvaardigheden zoals kennis van het alfabet en fonologische bewustzijn. Uit de derde meta-analyse bleek dat ongeveer 7% van de woordenschatgroei van zowel peuters als kleuters verklaard werd door interactief voorlezen op school. Bovendien bleken alleen de kleuters alfabetkennis op te doen; jongere peuters hadden kennelijk al hun aandacht nodig voor het begrijpen van het verhaal en letten nauwelijks op de geschreven tekst in prentenboeken. Onderzoekers bleken echter veel effectiever in het stimuleren van de woordenschat dan de eigen leerkrachten, die alleen bij het voorlezen aan de hele klas gemiddeld effectief bleken in het uitbreiden van de woordenschat van risicogroepkinderen. Het lijkt daarom nog te vroeg om een grootschalige implementatie van interactief voorlezen binnen het schoolcurriculum aan te bevelen. Om de effectiviteit van een interactieve voorleesstijl op school te vergroten, is het waarschijnlijk noodzakelijk om leerkrachten individueel te coachen en ze meer informatie te geven over de theorie achter de interventie.

Samenvattend kan gesteld worden dat de meta-analyses in dit proefschrift ondersteunen dat het lezen van (prenten)boeken een belangrijke bijdrage levert aan schoolsucces. Voorlezen lijkt onderdeel te zijn van een continuüm van leeservaringen, die steeds meer buiten schooltijd worden opgedaan. Deze leeservaringen faciliteren en beïnvloeden de taal-, lees- en spelvaardigheden door de hele ontwikkeling heen. Het is daarom belangrijk om te investeren in de kwaliteit van voorlezen aan jonge kinderen, zeker in gezinnen waar weinig ervaring is met boeken en in peuterspeelzalen en kleuterklassen van kinderen die met een verhoogd risico op taal- en leesachterstanden op school komen. De resultaten van twee meta-analyses toonden echter aan dat de effecten van interactieve voorleesprogramma's juist tegenvielen in risicogezinnen en in natuurlijke schoolsituaties van risicogroepkinderen. Voor vervolgonderzoek is het van belang om uit te zoeken hoe de kwaliteit én de kwantiteit van leeservaringen kan worden gepromoot bij zowel jonge kinderen als zwakke en goede lezers. Daarnaast is het nodig om kinderen longitudinaal te volgen binnen hun thuis- en schoolomgeving. Zo kunnen de processen en strategieën in kaart worden gebracht die verklaren hoe voorlezen aan kleine kinderen het leesgedrag van adolescenten en volwassenen bepaalt.

Dankwoord (Acknowledgements)

Mijn proefschrift had niet deze vorm gekregen zonder de kennis, vrijheid en belangstelling van mijn omgeving. Ik dank mijn collega's voor het samen werken, voor de interessante discussies over artikelen, statistiek en documentaires en voor de (trein)reizen. Ik had de "meta-a*mol*ysis" (Smeets, 2008) evenmin kunnen uitvoeren wanneer (inter)nationale wetenschappers hun onderzoeksresultaten niet beschikbaar hadden gesteld via tijdschriften, boeken, e-mail of het internet. Vriend(inn)en van de allereerste onderwijsgroep, van GGK en buiten GW, vriendinnen uit Houten, die soms ook naar Maastricht gekomen zijn, nu zelfs vlakbij wonen of die ik in Den Haag heb ontmoet: dank jullie wel voor jullie meeleven met mij en mijn dromen. Mijn familie dank ik voor de vaste thuisbasis en de liefde voor literatuur die jullie hebben doorgegeven. Egbert, dat ik jou in de laatste, dynamische maanden heb mogen leren kennen en dat we onze promotieperikelen konden delen, vind ik heel bijzonder.

Curriculum Vitae

Suzanne Mol was born July 30, 1983 in Naarden, the Netherlands. She completed her secondary education in 2001 (Gymnasium, Koningin Wilhelmina College in Culemborg). She studied Health Sciences at Maastricht University, specializing in Mental Health Sciences (cum laude, 2006). Her master's thesis focused on the intergenerational transmission of life aspirations in adolescents. During her studies, she was student-assistant of Dr. Susan Bögels (Department of Clinical, Medical, and Experimental Psychology) for more than two years; a period in which she also briefly participated in research of the Babylab at the University of Reading (U.K.). In December 2006, she started a PhD project in the Faculty of Social and Behavioural Sciences (Education and Child Studies; Leiden University). Her main aim was to conduct meta-analyses that quantitatively synthesize existing research in the field of (shared) book reading.